Vol. 40, No. 2 April, 1950

Journal of Experimental Psychology

EDITED BY

FRANCIS W. IRWIN

UNIVERSITY OF PENNSYLVANIA

CONSULTING EDITORS

CLARENCE H. GRAHAM, *Columbia University*
HARRY HELSON, *Brooklyn College*
DONALD B. LINDSLEY, *Northwestern University*
DONALD G. MARQUIS, *University of Michigan*
GEORGE A. MILLER, *Harvard University*
L. L. THURSTONE, *University of Chicago*
G. R. WENDT, *University of Rochester*
JOSEPH ZUBIN, *N. Y. Psychiatric Institute, Columbia University*

CONTENTS

The Effects of Experimentally Induced Needs upon Retention: M. WILLIAMS 139
Acquisition of a Simple Spatial Discrimination as a Function of the Amount of Reinforcement: B. REYNOLDS 152
Drive Specificity and Learning: The Acquisition of a Spatial Response to Food under Conditions of Water Deprivation and Food Satiation: E. L. WALKER, MARGARET C. KNOTTER, AND R. L. DEVALOIS 161
Studies in Motivation and Learning: II. Thirsty Rats Trained in a Maze with Food but No Water; Then Run Hungry: H. GLEITMAN 169
Acquired Distinctiveness of Cues: II. Selective Association in a Constant Stimulus Situation: D. H. LAWRENCE 175
Generalization of Secondary Reinforcement: I. J. SALTZMAN 189
A Preliminary Determination of the Functional Relationship of Effective Reaction Potential ($_{s}\bar{E}_{R}$) to the Ordinal Number of Vincentized Extinction Reactions (n): H. C. WILCOXON, RUTH HAYS, AND C. L. HULL 194
Effects of Competing Reactions on the Conditioning Curve for Bar Pressing: W. K. ESTES 200
The Effect of Perception Time upon Rigidity and Concreteness of Thinking: M. ROKEACH 206
Analysis of Detour Behavior: IV. Congruent and Incongruent Detour Behavior in Cats: P. H. SCHILLER 217
Learning Motivated by a Manipulation Drive: H. F. HARLOW, MARGARET K. HARLOW, AND D. K. MEYER 228
The Relationship between Two Measures of Response Strength: H. SCHLOSBERG AND C. HEINEMAN 235
Evidence for the Role of Motivation in Determining the Amount of Reminiscence in Pursuit Rotor Learning: G. A. KIMBLE 248
The Role of Recency in Learning: R. H. WATERS AND J. G. REITZ 254
The Relation of Structural Changes of the Eye and Vitamin A to Elevation of the Light Threshold in Later Life: J. E. BIRREN, M. W. BICK, AND M. YIENGST 260
The Destruction of the Müller-Lyer Illusion in Repeated Trials: I. An Examination of Two Theories: W. KÖHLER AND JULIA FISHBACK 267
A Comparison of Critical Flicker Frequencies under Conditions of Monocular and Binocular Stimulation: F. H. IRELAND 282

PUBLISHED BI-MONTHLY BY
THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
PRINCE AND LEMON STS., LANCASTER, PA.
AND 1515 MASSACHUSETTS AVE., N. W., WASHINGTON 5, D. C.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879
Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published bi-monthly by the American Psychological Association, Inc. The annual subscription rate is $7.00 (foreign, $7.50); single copies, $1.25. Subscriptions, orders, and business communications should be addressed to the American Psychological Association, Inc., 1515 Massachusetts Ave., N.W., Washington 5, D. C.

This JOURNAL publishes articles which are intended to contribute toward the development of psychology as an experimental science. Experimental work with normal human subjects is favored over work with abnormal or animal subjects. Studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental psychology or for psychological theory. Articles of a length exceeding 25 printed pages cannot be accepted.

With rare exceptions, articles are published in the order of their receipt. Authors are supplied with 50 free offprints without covers. All of the cost of author's alterations in galley proof is charged to the author, and one-half of the cost of tables, cuts, and formulas.

Immediate publication can be given to articles whose authors assume the full cost of publication. Such articles, if submitted before October 1, 1950, should be addressed to the retiring editor, Francis W. Irwin, Department of Psychology, University of Pennsylvania, Philadelphia 4, Pa. No free reprints are given in the case of these articles.

Address articles submitted for regular publication, or for immediate publication after the December, 1950 issue, to the editor-elect, Arthur W. Melton, Human Resources Research Center, Lackland Air Force Base, San Antonio, Texas.

This JOURNAL adheres to the conventions concerning bibliographies, citations, preparation of tables and cuts, etc., as described by J. E. Anderson and W. L. Valentine, The preparation of articles for publication in the journals of the American Psychological Association, *Psychol. Bull.*, 1944, 41, 345–376. Special attention is called to the style of numbers and abbreviations (p. 353). The term 'experimenter,' with its various forms, singular, plural, and possessive singular and plural, is abbreviated *E*, *Es*, *E*'s, and *Es*'; analogous abbreviations are used for the terms 'subject' (*S*) and 'observer' (*O*). All manuscripts must be typed double-spaced throughout, *including tables, footnotes, and references.* Tables, figures, and footnotes must be placed on separate pages at the end of the manuscript.

Journal of
Experimental Psychology

VOL. 40, NO. 2 APRIL, 1950

THE EFFECTS OF EXPERIMENTALLY INDUCED NEEDS UPON RETENTION [1]

BY MEYER WILLIAMS

Veterans Hospital, Downey, Illinois

INTRODUCTION AND PROBLEM

Dynamic or motivational approaches to personality have long maintained that memory is an active selective process. Selection is said to be directed, to a large extent, by motivational forces or conditions of the individual which direct, organize, and define the adequacy of his behavior. It has been assumed that these motivational forces consciously or unconsciously influence what will be retained. Material is said to be differentially retained in accordance with whether or not it is related to these motivational forces. Material related to the needs, drives, attitudes, etc. will be retained better than non-related material.

Few experimental studies have attempted to evaluate the selective effects of specific needs upon retention. Laboratory psychology has emphasized investigation of stimulus dimensions while generally ignoring subject differences, particularly those of a motivational nature. Yet, retention as we observe it in everyday life does not appear to be the relatively passive, seemingly unselective performance common to most laboratory experiments. In the related area of perception, numerous studies (3, 7, 9, 11, 14, 16) have demonstrated that perception is preferential or selective and that selective perception results in part from the needs of the individual at the time of perceiving. If, as Murphy (10) points out, memory is a continuation or reactivation of central processes intimately related to perception, we might expect the same motivational influence to apply in retention as in perception. A few related experimental studies are to be found which claim that various motivational factors are extremely important in influencing what will be retained. These include studies by Barrett (1) on the influence of mental set, Fox (5) on the selectivity of subjective preference, Bartlett (2) on personal attitudes and Seeleman (18), Levine and Murphy (8), Edwards (4), and Postman (13) on racial, political and social attitudes. Superficially, the results of all the above retention studies appear to be in accord; namely, selective motivational effects appear to be influential in accentuating the retention of motivationally related material. However, more careful consideration of the above retention studies calls their findings of selective retention into question.

First, in each of the above studies it is obvious that the specific motivational state was present in the subject at the time of the original learning of the related material. We might thus expect

[1] The writer wishes to express his appreciation to Dr. Benton J. Underwood for his valuable criticism and aid during the development and execution of this study.

differential learning favoring need or attitude-related material to have occurred and thereby account for the obtained differences in retention. That this was the case is borne out by a number of the studies cited in which such differences were found during original learning. We would therefore expect better retention for need or attitude-related material since such autistic material was learned better originally.

Secondly, with the possible exception of the Barrett experiment, the aforementioned studies are relatively limited in their control of, and in their knowledge of the specific motivational factors which were primary in influencing differential retention. Even the Barrett study, in which set was controlled, appears vague as to the specific motivational factors involved. Use of subjective ratings or questionnaires gives little indication as to the relative validity, degree of strength, or origin of the selective processes at work. Until such motivational variables are subjected to experimental control, our basic understanding of selective retention is hampered by a vagueness as to the contributory factors in the motivational framework.

The above criticisms were not made to depreciate the importance of the foregoing pioneer studies. Rather they were made to suggest that our knowledge as to what motivational processes are really operating in selective retention is as yet superficial. For an adequate test of selective retention, two criteria must be applied: (1) the degree of original learning must be fully controlled and unbiased by the presence of specific motivational states related to the learning material, and (2) motivational factors must be adequately defined and controlled and, most important, must be introduced after original learning. It is with these criteria in mind that a series of experiments was designed to test critically the hypothesis of selective retention of motivationally related material.

The Problem.—The basic experimental rationale may be indicated by considering the differential treatment of two subjects. Two matched subjects learn a list of materials at the same rate and under the same standardized conditions. One-half of this list of materials relates to a particular need condition; the other half is neutral. This need condition is now experimentally induced in one of the subjects *after original learning.* At the same interval after original learning, both subjects are measured for retention of the list of materials. What should the results be? According to dynamic theory as well as the consensus of the past reviewed experiments, it may be expected that there will be better retention of the need-related portion of the material by the subject in whom the need was induced. Fundamentally, it is this hypothesis, with some extension, that the present experiments have attempted to evaluate.

The experiments which are outlined in the next section have investigated the effects of experimentally induced needs on two levels, (1) the biological, and (2) the psychological. In addition, study was attempted of a variation of this hypothesis, namely, that neutral material which becomes related to need-related material through association will also show selective retention in its favor. Thus the hypothesis was: The presence of a strong need, whether biological or psychological in nature, will bring about a differential selectivity in retention favoring: (a) material related to that need, and (b) neutral material related to that need through association with need-related material.

Procedure

Experiment I

General Procedure.—Experiment I was designed to determine the effects of an experimentally induced need for food upon the retention of

material related to that need. It employed a sequence of four sessions spaced 24 hours apart as follows:

Day 1—Subjects practiced the learning of a list of neutral material formally similar to that to be used in the experimental list.

Day 2—Practice on the recall and relearning of the Day 1 list was given. This was followed by the learning of another similar neutral list. Subjects were tentatively paired on the basis of their acquisition performance on this second list. Following this pairing, members of each pair were randomly assigned to either experimental or control group.

Day 3—Featured the learning of the experimental list consisting of 16 paired associates in which half of the response words were related to food or eating. In addition, a five card Thematic Apperception Test was administered after learning in order to mask and minimize the importance of the learning material and thus reduce the possibility of rehearsal. All subjects assigned to the experimental group were then placed on a 24-hour fast. All control subjects ate normally.

Day 4—Subjects of both groups were measured for retention of the experimental list.

Subjects.—Experiment I employed a total of 20 males ranging in age from 18 to 25. As indicated above, this group was randomly divided into experimental and control groups of 10 *Ss* each. All *Ss* were volunteers who had been informed previously that the experiment might involve a hardship or stress condition. Prior to volunteering they had also been told that participation in the experiment would entitle them to twice the experimental credit usually given in Psychology A classes. As a consequence, motivation to participate and fulfill the experimental requirements appeared high, and the attrition rate was negligible.

Learning Materials.—An experimental list of 16 paired associates of approximately zero degree association value was selected on the basis of the Kent-Rosanoff Free Association Norms (15). (Zero degree association value was defined as that degree of association between a stimulus word and a response word characterized by the fact that less than two persons in 1000 gave the response word as an association to the stimulus word.) Stimulus words were one-syllable adjectives taken from the Kent-Rosanoff Stimulus Word List. Response words were nouns of one syllable. Half of these response words were names of common foods. The other half were words neutral in meaning in relation to the topic of food or eating. All stimulus words were likewise neutral and were matched to response words in such a manner that phonetic, formal structural, or other special clues were minimal. There thus resulted a list of 16 paired-associates which could be visualized as being made up of two sub-lists, one neutral in terms of the need for food, the other directly related to the need for food through its response words. These are presented below:

cold—bird; swift—tree; soft—room; high—book; green—cliff; light—road; loud—height; long—stone.

dark—food; square—cheese; slow—bread; blue—steak; short—beef; white—pear; deep—milk; hard—cake.

The lists which were utilized during the two practice days were identical in length, association value, and formal similarity to the experimental list. Neutral words were used throughout in both practice lists.

Learning Procedure.—The paired-associates lists were presented by means of a Hull-type memory drum. In order to hold serial effects to a minimum, three different list orders were used in a standard sequence. In this manner each paired associate appeared in a completely different list position (beginning, middle or end of list) during each succeeding presentation. Learning of the paired associates was by the anticipation method utilizing a rate of one pair per 4 seconds. The criterion of learning was one perfect recitation of the list. Instructions to *Ss* were those commonly used for paired associate learning.

Induction of Need.—Induction of the need for food in the experimental group was accomplished by strict instruction to the *Ss* to begin fasting for the next 24 hours. This instruction was given immediately following the above learning and TAT session. The importance to the experiment of complete asbtinence from food was stressed at this time. *Ss* were allowed to take only water during the period of fasting. All *Ss* (control and experimental) used in this experiment underwent their learning and TAT session (Day 3) shortly after a meal (zero to three hours), and thus may be considered to have been relatively satiated at the time of learning. A nonpunitive interview was conducted with each experimental subject at the end of his Day-4 session in which he was made to feel free to report any violations of fasting without fear of recrimination or loss of experimental credits.

Retention Procedure.—Recall and relearning of the experimental list were utilized as measures of retention for all groups 24 hours after learning. The sequence and manner of presentation of materials during the retention session was the same as during the learning sessions. However, each subject began his trials with the list order which would have ordinarily followed his last learning trial. Instructions to the subject dur-

ing this session were: "I am now going to present the list of paired words which you learned during the last session. Your instructions are the same as before. Be sure you give as many correct responses as you can on the trial and throughout the relearning. You will be through with the list when you can call out correctly the second word of every pair for an entire presentation."

Experiment II

General Procedure.—Experiment II was designed to determine the effects of an experimentally induced need for food upon the retention of neutral material which had become related to the food need through association with food material.

Experiment II duplicated Experiment I in every detail with the exception that the experimental list was reversed, the adjectives being used as response words and the nouns as stimulus words. The experimental paired-associates list thus contained one-half food and one-half neutral items as stimulus words while all response items were neutral in character. The Rosenzweig Picture-Frustration Test was used instead of the TAT as a masking device in this study, and was administered following learning of the experimental list.

Experiment III

General Procedure.—Experiment III was designed to determine the effects of an experimentally induced need for release of hostility upon the retention of material related to that need. As in the first two experiments, experimental and control groups of 10 *S*s each were employed, *S*s being tentatively paired on practice-learning results. However, the experimental sessions were conducted over three consecutive days, 24 hours apart, instead of over four days in the previous studies.

Day 1—*S*s learned the same practice list as that given on Day 2 in Experiments I and II (Practice List I was omitted).

Day 2—Retention of this practice list was measured, and the list was relearned. This was followed immediately by the learning of the experimental list. Upon the conclusion of the learning of this experimental list, a sentence completion test was administered to provide the masking effects as in the previous experiments. (The sentence completion test is also a projective device made up of a group of incomplete sentences. The *S* is instructed to give the first completion that comes to mind for each sentence.)

Day 3—*S*s in the experimental group underwent a social stress situation in which they received intense personal criticism and frustration. Immediately following this situation, *S*s of the experimental group were examined for retention of the experimental list in the same manner as described in Experiment I. *S*s of the control group were also examined for retention of the list after the same lapse of time as the experimental group. The experimental list was identical to that in Experiment I except that words related to hostility were substituted for the words related to food and eating.

Subjects.—Twenty *S*s were selected on the same basis as those in Experiment I.

Learning Materials.—Experimental learning materials were the same list as that used in Experiment I, except that eight response words related to aggressive and hostile activity which could be either directed outward or toward the self were substituted for the eight hunger response words. The same precautions as to association value and formal similarity were applied as in Experiment I. The experimental list was as follows:

cold—bird; swift—tree; soft—room; high—book; green—cliff; light—road; loud—height; long—stone.

dark—blow; square—harm; slow—fight; blue—rage; short—hate; white—threat; deep—curse; hard—death.

Learning Procedure.—Learning procedure was identical to that in Experiment I.

Induction of Need.—Twenty-four hours after the experimental learning and TAT session, each *S* in the experimental group was admitted to the viewing room of a one-way screen situation (the same room in which learning took place) and was introduced to a 'standard group of psychologists' sitting about the exposed one-way screen. A motion picture camera was placed prominently against the screen. The operation of the one-way screen was demonstrated to the *S*, following which the *S* was seated at a table flush against the opposite or mirror side of the screen in the observation room. The observation room was dark except for two photo-flood lights focused on the *S* and the area before him. The table at which he was sitting was bare except for a pad of test blanks and some pencils. At this point, the experimenter informed the *S* that the next part of the experiment would take place in this room and that all further instructions would come via a concealed speaker. An elastic electrode band was then placed on the *S*'s non-writing hand. The experimenter re-entered the viewing room closing the door between the two rooms and leaving the *S* alone in the observation room.

The following standard text and procedure were employed utilizing a low-powered P.A. system for communication with the *S*:

"Mr. ———, as you know although you cannot see them you are now being observed by a number of psychologists on the other side of the screen who are taking notes and motion pictures of all your reactions throughout this experimental session. All directions are to be followed quickly and with no comment. Save any questions for the end of the experiment. Now listen carefully—pick up a pencil and look at the top of the test sheet in front of you. You will notice that the top line contains a series of numbers each of which has a figure or symbol directly under it. That figure or symbol stands for the number above it. Below this group or code if you wish, you will notice a long series of numbers which have no figures or symbols in the blank spaces under them. Your task is to draw the proper figure under each of these numbers using the code at the top of the page. Let's practice now—what figure goes under the first number, the second, the third . . . O.K. Be sure your figures are accurate. Now finish the line . . . O.K. Give me your complete attention. Your goal in this test is to finish all the items on the test sheet from the time I give you a "Go" signal until you are told to stop. Do each item in order and do not stop to correct items.

"As you may have gathered, this test is one frequently used in intelligence measurement and gives a pretty good index to your mental ability. You will have a number of trials in which to demonstrate your ability. Most people of normal intelligence are able to finish the entire test sheet in the standard test time, so being a college student you will undoubtedly do well. However, just to make sure you are thoroughly motivated, the psychologists out here will appraise your efforts at the end of each trial and if you do not meet their standards, you will be shocked—like this . . . (Subject received standard one second duration shock of moderate intensity.) Pay attention now—let's see how fast you can finish the remaining items on the first page—Get ready—Go! (The subject was allowed to finish the page before time was called.) Fine! You did very well for a first try. Just keep it up—you should improve a great deal with practice. Watch the code carefully as it will be different during each trial. Turn to the next page. In a few moments you will receive the "Go" signal for the next trial. Remember, your reactions are being photographed and observed by the psychologists on the other side of the screen, and if you do not meet the standards you will be shocked. . . . Get ready for trial two. Go! (A standard 60-sec. period was allowed for each trial from this point on.) Stop! (Standard shock was administered.) You should have been able to finish that page. Perhaps you will improve next time—remember, your actions are being observed and photographed and you will be shocked if you do not meet our standards. Turn to the next page and get ready for trial three. Go! (60-sec. trial) Stop! (Standard shock administered.) That was a very poor performance. You certainly can do better than that. Any person of normal intelligence would have finished that page. You certainly can't be trying too hard. We'll give you another chance. Turn the page and get ready for trial four. Go! (60-sec. trial period) Stop! (Electric shock administered.) I guess you just can't improve. Very few people are unable to finish that page by the fourth trial. I guess you just don't have what it takes for this test. We might as well quit. Remain seated until the apparatus is removed. Maybe you'll do better in the next part of the experiment."

Retention Procedure.—Retention procedure was identical to that used in Experiment I. Control subjects were also given the digit symbol test sequence prior to the measure of retention of the experimental list. However, this sequence was given under normal rather than social frustration conditions, practice and test trials being identical in length (60 sec.) as were intervals between trials (30 sec.).

Experiment IV

General Procedure.—Experiment IV was designed to determine the effects of an experimentally induced need for release of hostility upon the retention of neutral material which had become related to the need through association with material with hostile connotations.

Experiment IV duplicated Experiment III throughout, with the exception that, as in the case of Experiment II, the experimental list was reversed. The experimental paired-associates list thus contained one-half hostile and one-half neutral items as stimulus words, while all response items were neutral or innocuous.

The Rosenzweig Picture Frustration Test was employed as the masking device in this study at the conclusion of the learning of the experimental list.

Results

Acquisition.—In each of the experiments, evaluation of existing differences in retention between experimental (E) and control (C) groups is contingent upon comparable acquisition levels on the experimental list. It is thus necessary to demonstrate that the original learning levels for

TABLE I

A Comparison of Original Levels of Learning Attained by Experimental (E) and Control (C) Groups on Respective Experimental Lists (N = 10 in each group)

Experiment	E Mean	E σ_m	C Mean	C σ_m	Mean Diff.	$\sigma_{m.d.}$	t*
I	19.7	2.80	21.1	2.79	1.4	3.95	.35
II	13.1	2.04	11.6	1.82	1.5	2.73	.55
III	24.5	4.04	25.9	3.12	1.4	5.10	.28
IV	15.5	1.97	15.7	2.25	.2	2.99	.07

* t-values required at the 1 and 5 percent levels of confidence are 2.88 and 2.10 respectively.

each pair of E and C groups is substantially the same. The essential data are presented in Table I in terms of the mean number of trials required to attain a perfect recitation of the experimental list.

For all experiments it will be noted that t-values are clearly insignificant, indicating that in each case the performances of E and C groups were relatively equal. Since no reliable differences exist in original learning, differences found on the retention measures must be interpreted as a product of the induced need conditions.

Retention.—In order to introduce even greater precision into the evaluation of the retention results, E and C S's of each experiment were paired by scores based upon acquisition rate of the experimental list. (Pairing on the acquisition rates of need-related materials alone was also carried out. The results obtained were of the same order as those found by pairing on total acquisition rate.) Differences in learning were thus held to a minimum. Since the major tests of the selective retention hypothesis lie in the results related to the retention of critical or need-related material, the retention data were analyzed into two sections, one dealing with results on need-related material and material associated with need related material, and the other dealing with purely neutral material.

Retention of need-related material and material associated with need-related material is considered first. Four measures of retention were employed in evaluating this material: (1) first recall, (2) second recall, (3) trials to relearn to the criterion, (4) number of items missed in relearning (no responses plus erroneous responses). The essential retention data utilizing the above measures are shown in Table II.

According to the selective retention hypothesis, expected results would be as follows:

1. First Recall—E greater than C
2. Second Recall—E greater than C
3. Trials to Relearn—E less than C
4. Items Missed in Relearning—E less than C

Observed results on First Recall show a consistent trend in the expected direction but t-values do not approach significance at any reliable level.

With the exception of Experiment III, Second Recall indicates no significant differences or consistent trends in favor of the hypothesis. Results of Experiment III are in the expected direction and are reliable at the five percent level. However, since this is the only significant t-value appearing in all the retention measures, its importance would seem minimal.

The results on Trials to Relearn are likewise statistically insignificant and inconsistent. Again Experiment III and also IV show results in the expected direction but the differences are not great enough to be reliable.

On Number of Items Missed in Relearning, Experiment III once more presents a relatively pronounced though still unreliable trend in the expected direction. Differences on the other three experiments are too small to be of any significance.

In general, then, the results indicate no significant differences favoring the retention of need-related material or neutral material associated with need-related material. A consistent but unreliable trend appears on the first recall, in which E groups show a slightly greater tendency to recall need-related material than do C groups. In addition, the E group of Experiment III (induced need for hostile expression) shows a consistent but unreliable trend toward better retention of need-related material on all measures than does its C group. In summary, the results do not offer any reliable evidence in favor of the

TABLE II

A Comparison of Experimental (E) and Control (C) Groups on Various Measures of Retention of Need-Related Material (N = 10 for each group)

Experiment	E Mean	E σ_m	C Mean	C σ_m	Mean Diff.	$\sigma_{m.d.}$	t*
			First Recall				
I	5.7	.26	5.4	.31	.3	.37	.82
II	6.1	.43	5.4	.31	.7	.54	1.30
III	5.8	.39	4.8	.57	1.0	.86	1.17
IV	5.9	.43	5.4	.34	.5	.40	1.25
			Second Recall				
I	6.7	.42	7.2	.33	.5	.43	1.17
II	7.1	.28	7.3	.18	.2	.29	.69
III	7.0	.26	5.9	.38	1.1	.43	2.54
IV	6.9	.38	6.8	.29	.1	.53	.19
			Trials to Relearn				
I	4.1	1.05	3.2	.53	.9	1.24	.73
II	3.3	.50	3.0	.30	.3	.54	.56
III	4.2	.47	5.5**	1.09**	1.3	1.02	1.27
IV	3.4	.65	4.2	.85	.8	.74	1.08
			Items Missed in Relearning				
I	5.9	1.54	5.3	1.09	.6	1.61	.37
II	3.8	.93	3.9	.62	.1	1.09	.09
III	6.1	.85	11.2**	2.74**	5.1	2.71	1.88
IV	5.4	1.59	6.0	1.61	.6	1.95	.31

* t-values given in this table are based upon differences involving paired observations. The t values required at the 1 and 5 percent levels of confidence are 3.25 and 2.26 respectively.

** These comparatively inflated values are due to a single S in the control group whose scores on these retention measures departed radically from those of the rest of his group.

TABLE III

A Comparison of Experimental (E) and Control (C) Groups on Various Measures of Retention of Neutral Materials (N = 10 for each group)

Experiment	E Mean	E σ_m	C Mean	C σ_m	Mean Diff.	$\sigma_{m.d.}$	t*
First Recall							
I	5.8	.36	5.1	.35	.7	.56	1.25
II	5.3	.58	5.3	.40	.0	.70	.00
III	6.6	.43	6.6	.45	.0	.60	.00
IV	6.0	.38	6.5	.48	.5	.65	.77
Second Recall							
I	7.2	.33	6.9	.24	.3	.21	1.41
II	7.0	.37	7.0	.21	.0	.45	.00
III	7.0	.26	6.9	.50	.1	.64	.16
IV	7.3	.21	6.9	.35	.4	.40	1.00
Trials to Relearn							
I	3.8	.84	3.8	.59	.0	1.11	.00
II	2.9	.57	3.4	.40	.5	.65	.77
III	2.7	.42	3.4**	1.09**	.7	1.17	.60
IV	3.3	.56	3.1	.82	.2	1.01	.20
Items Missed in Relearning							
I	5.8	1.87	6.0	1.12	.2	2.19	.09
II	4.6	1.13	4.5	.79	.1	1.35	.07
III	3.8	.85	6.5**	2.29**	2.7	1.97	1.37
IV	3.8	1.05	4.2	1.71	.6	2.14	.28

* t-values given in this table are based upon differences involving paired observations. The t-values required at the 1 and 5 percent levels of confidence are 3.25 and 2.26 respectively.

** These comparatively inflated values are due to a single S in the control group whose scores on these retention measures departed radically from those of the rest of his group.

selective retention hypothesis, nor do they support the findings of previous related studies.

Retention of neutral items was analyzed by means of the same measures employed for need-related items. Essential data for neutral material are presented in Table III.

The results on the retention of neutral items need little discussion. As may be seen, differences approaching reliable confidence levels are not in evidence. Inspection of the data reveals no consistent direction of the differences either for a particular retention measure or for a given experiment. It may be concluded that retention of neutral material appears uninfluenced by (unrelated) need conditions.

Discussion of Results

General Interpretation of Major Results.—The conclusions which may be drawn from the major results of this study appear clear cut. In critical tests in which original learning was relatively free from related needs and in which related needs were induced after original learning:

1. Retention of need-related material was not reliably influenced.

2. Neutral material associated with need-related material during original learning was not differentially retained.

3. Retention of neutral material, not directly related to the induced needs, was unaffected by the presence of the needs.

4. Biological need (hunger) did not show any reliable difference from psychological need (release of hostility) in bringing about differential retention.

In general, then, the importance of need in the retention of need-related or neutral material is nil or, at most, of minor importance (if the consistent but unreliable trend on first recall is given weight). If these conclusions are extended, past related studies claiming to demonstrate selective retention have, in reality, demonstrated differences in retention as a function of motivationally induced differences in original learning. Differential retention in these past studies was thus apparently a product of selective learning and not of some selective process taking place after learning.

Theoretical Interpretation of Major Results.—Theoretical interpretation of the major results offers a difficult problem. Modern theories of learning have, as yet, not advanced to the point of providing other than general statements about the relation of motivation to retention. Personality theories arguing for selective retention or some motivational-enhancement process after original learning are not supported by the results. It is therefore necessary to construct tentative theoretical explanations which will take into consideration both past and present empirical findings and which can be subjected to experimentation.

Three possible theoretical explanations present themselves. The failure to find selective retention may be due to the fact that:

1. Selective retention does not exist; apparent selective retention observed in everyday life and in clinical practice is due to differences in original learning or to invalid observations.

2. The memory drum situation is too highly structured and too unlike the retention situations of everyday life to allow a complex and dynamic process such as selective retention to exhibit itself.

3. Interaction between the induced needs and the need-related portions of the learned lists did not occur since the set to perform in the memory drum situation was dissimilar or incompatible with the set produced by the need.

The first interpretation, that apparent selective retention is simply a function of differences in original learning, appears most logical in the light of the present results. However, since there is a paucity of related studies, it is difficult to estimate how far one can generalize from the present results. For instance, it is not known whether studies utilizing different or more intense needs, lower degrees of original learning, longer rest intervals between learning and retention, different types of verbal materials, and different learning and retention situations would show results identical to those of the present study. Until further related studies such as suggested above are carried out, the interpretation that selective retention does not exist must be tentative. At least, such a conclusion must not preclude the possibility of other interpretations of the present results.

The second explanation, which suggests that the absence of selective retention is a function of a high degree of structuring in the memory drum situation, is one which can hardly be overlooked. The restrictions imposed by highly controlled laboratory procedures with humans constitute a problem which experimental psychology has yet to meet adequately. Employment of a memory drum situation such as was used in the

present study is valuable in terms of the relatively rigid control it provides over pertinent stimulus dimensions. On the other hand, the memory drum situation hardly approximates the manner in which most individuals learn or recall in everyday life. Retention situations of everyday life are characterized by freedom to sift, accept and reject material. Furthermore, retention of ideas rather than of specific isolated words is important. Undoubtedly, the free everyday retention situation also allows much greater ineraction between stimulus variables and intervening variables such as needs. If selective retention is to exhibit itself in experimental studies, the likelihood of its appearance may well be a function of this freedom to integrate present need variables with past learning variables. Such freedom is obviously curtailed in the highly controlled memory drum situation and this may possibly account for the failure to find selective retention.

The third interpretation, that of no interaction between dissimilar sets, is related to the second interpretation in that it also argues that the failure to find selective retention resulted from lack of interaction between the need and the need-related material. However, in this explanation, the absence of selective retention is interpreted to be due to differences in set rather than to a restricted retention situation. (In view of the numerous meanings attributed to the term *set* (6), it may be well to define *set* as it is used in the present discussion. *Set* refers to a learned predisposition to react in a fixed manner to a given stimulus situation.) Perhaps this interpretation is the most promising of the three since it is supported by the results of the experimenter's interviews with subjects and by other studies.

The theoretical explanation takes the following form. Subjects learned the experimental lists to a high degree. At the time of learning they had no inkling that needs were to be induced which were related to half of the material. Learning the list meant simply adopting a familiar academic-like set toward reaching a prescribed academic-like goal. The entire list was thus learned as an isolated group of materials relatively unrelated to any other group. Induction of the need after original learning was effective. However, little perception of the relationship between the need and the need-related half of the material occurred since the set induced by the need was dissimilar from or incompatible with the academic-like set of the memory drum situation. Furthermore, since the measure of retention took place relatively soon after original learning, the effects of original learning were still extremely strong. During retention, the original memory drum or learning set was reinstated and the need set was temporarily inhibited. Retention was simply in accordance with the effects of original learning. As a consequence, no selective retention related to need was in evidence.

This explanation is supported by interviews. When subjects began discussing the experiment with the experimenter after their final session, the majority appeared unaware of the relationship of the need-related materials to the need condition. However, the moment the experimenter asked the question, "Did you notice anything in common between the words you learned and the experimental condition?" most subjects expressed sudden insight into the relationship and were astonished that they had not recognized it before.

This interpretation is further supported by studies in other areas. Sears, Hovland, and Miller (17) conducted a study in which frustration was induced by keeping subjects awake throughout the night. In addition, they actually broke previously made promises and forced subjects to undergo irritating tasks during this period. Time-sample observation during this sleepless period showed marked increases in aggressive speech and behavior. Spontaneous aggression drawings were produced and were overtly identified with hostility toward the experimenters. To measure further the aggression, the experimenters administered five paper-pencil tests dealing with response dimensions expected to change with frustration. In spite of all

the overt evidence of aggression, results on the five paper-pencil tests showed no significant differences from control group results.

The conclusion reached by these experimenters was that paper-pencil tests involve a well-established set to carry out instructions. The generalized aggressive set caused by the frustration was new and relatively momentary in character. In addition, it was incompatible with the paper-pencil task set. When confronted with the paper-pencil tests, subjects responded in terms of the well-established task set and temporarily inhibited the new aggressive set. As a result, performance on the paper-pencil tasks was in line with the manner in which the subject ordinarily responded to such tests.

In an experimental evaluation of the judgmental theory of feeling, Peters (12) likewise has shown that an old judgmental set may reassert itself despite an interpolated differential re-inforcing activity. He presented subjects with a series of colored pictures and instructed them to judge these pictures on a pleasantness-unpleasantness dimension. Later he had these subjects utilize a portion of these pictures in successfully solving a learning problem. This was calculated to enhance the pleasantness of these pictures. At still a later interval, the subjects were again asked to judge the entire original group of pictures along the pleasantness-unpleasantness dimension, and to indicate in writing the characteristic rationale for making pictorial judgments.

A control group also made judgments at an identical interval but without the interpolated learning problem. No significant changes in judgments were shown by the control group. However, the experimental group exhibited a dichotomy in judgmental results. One portion of the experimental group showed definite changes in pleasantness favoring the pictures involved in the problem; the other portion showed no change from original judgments. Change or lack of change was found to correlate highly with the subject's judgmental set as indicated in his written description. Subjects showing change indicated that they had used a 'concrete set' involving feelings about the actual qualities of the picture such as the colors, the color combinations, and the form and structure of the picture. On the other hand, subjects showing no change characteristically used an 'abstract set' involving associations and thoughts aroused by the pictures. Apparently no interaction took place between the original judgments and the successful learning of some of the material when subjects characteristically judged under an 'abstract set.'

Whether or not Peters' results lend weight to a judgmental theory of feeling is of no moment to the present discussion. What is important, however, is Peters' demonstration that an original 'abstract set' can withstand the effects of an intervening and dissimilar task set involving the original material. Moreover, it can be reinstated in its original form and with its original effect on this same material.

In summary, then, the explanation of the absence of selective retention in the present study in terms of dissimilar or incompatible sets is supported by findings in other areas. However, it is recognized that the evidence marshalled in the foregoing paragraphs to back up this explanation is at a lower confidence level than might be desired. This is a situation which exists throughout the area of interaction of motivational forces. It can only be remedied by further experimentation with pertinent dimensions.

Summary

Four experiments were performed to test critically the hypothesis that the presence of a strong biological or psychological need would bring about differential selectivity in retention favoring (a) material related to that need, and (b) neutral material related to that need through association with need-related material. Two criteria were employed for an adequate test of the hypothesis: (1) The degree of

original learning was fully controlled and unbiased by the presence of specific motivational states related to the learning material, and (2) Motivational factors were defined and controlled and, most important, introduced after original learning.

Experiment I was designed to test the hypothesis that a hunger need induced after learning would bring about selective retention of previously learned material related to hunger.

Experiment II was designed to test the hypothesis that a hunger need induced after learning would bring about selective retention of previously learned neutral material related to hunger through association with hunger material.

Experiment III was designed to test the hypothesis that a need for hostile expression induced after learning would bring about selective retention of previously learned material related to the expression of hostility.

Experiment IV was designed to test the hypothesis that a need for hostile expression induced after learning would bring about selective retention of previously learned neutral material related to the need through association with hostile material.

Conclusions based on the major results of all four experiments were that induction of a need after learning does *not* influence:

1. Retention of need related material.
2. Retention of neutral material previously associated with need related material.
3. Retention of neutral material not directly related in any manner to the need.

In addition, it was concluded that biological need was not reliably different from psychological need in influencing retention.

Three interpretations of the major results were offered and their implications discussed. It was suggested that the absence of selective retention in the present study might have been due to the fact that:

1. No selective retention exists after learning.
2. The memory drum situation was too highly structured and controlled.
3. Interaction between induced needs and need-related materials did not take place because of the dissimilarity or incompatibility existing between the academic-like task set and the need set.

(Manuscript received May 18, 1949)

References

1. Barrett, D. M. Memory in relation to hedonic tone. *Arch. Psychol.*, 1938, No. 223.
2. Bartlett, F. C. *Remembering, a study in experimental and social psychology.* Cambridge: Univ. Press, 1932, 206–207.
3. Bellak, L. The concept of projection: an experimental investigation and study of the problem. *Psychiatry*, 1944, 7, 353–370.
4. Edwards, A. L. Political frames of reference as a factor influencing recognition. *J. abnorm. soc. Psychol.*, 1941, 36, 34–50.
5. Fox, C. The influence of subjective preference on memory. *Brit. J. Psychol.*, 1923, 13, 398–404.
6. Gibson, J. J. A critical review of the concept of set in contemporary experimental psychology. *Psychol. Bull.*, 1941, 38, 781–817.
7. Levine, R., Chein, I., & Murphy, G. The relation of the intensity of a need to the amount of perceptual distortion: a preliminary report. *J. Psychol.*, 1942, 13, 283–293.
8. Levine, M. M., & Murphy, G. The learning and forgetting curve of controversial material. *J. abnorm. soc. Psychol.*, 1943, 38, 507–515.
9. McClelland, D. C., & Atkinson, J. W. The projective expression of needs: I. The effect of different intensities of the hunger drive on perception. *J. Psychol.*, 1948, 25, 205–222.

10. Murphy, G. *Personality: a bio-social approach to origins and structure.* New York: Harper, 1947, 372.

11. Murray, H. A. The effect of fear upon the estimates of maliciousness of other personalities. *J. soc. Psychol.*, 1933, **4**, 310–329.

12. Peters, H. N. Experimental studies of judgmental feeling: IV. Concrete versus abstract sets in the preference judgment of pictures. *J. exp. Psychol.*, 1943, **33**, 487–499.

13. Postman, L., & Murphy, G. The factor of attitude in associative memory. *J. exp. Psychol.*, 1943, **33**, 228–238.

14. Rodnick, E. H., & Klebanoff, S. G. Projective reactions to induced frustrations as a measure of social adjustment. *Psychol. Bull.*, 1942, **39**, 389.

15. Rosanoff, A. J. *Manual of psychiatry.* New York: Wiley, 1938.

16. Sanford, R. R. The effects of abstinence from food upon the imaginal processes: a further study. *J. Psychol.*, 1937, **3**, 148–159.

17. Sears, R. R., Hovland, C. I., & Miller, N. E. Minor studies in aggression: I. Measurement of aggressive behavior. *J. Psychol.*, 1940, **9**, 275–295.

18. Seeleman, V. The influence of attitude upon the remembering of pictorial material. *Arch. Psychol.*, 1939, No. 238.

ACQUISITION OF A SIMPLE SPATIAL DISCRIMINATION AS A FUNCTION OF THE AMOUNT OF REINFORCEMENT[1]

BY BRADLEY REYNOLDS

University of Missouri

Introduction

In two studies already reported (6, 7) the relationship between the strength of a habit and the amount of reinforcement has been investigated. One of these involved a black-white discrimination, the other instrumental conditioning. In neither was any relationship observed between the strength of a habit and the amount of reinforcement. The results of these studies were consistent with implications drawn from what has been designated the R hypothesis. In the study to be reported here we have examined certain other implications of this hypothesis.

The R hypothesis states that a positive relationship holds between the amount of reinforcement and certain measured aspects of a limited class of learned responses. These are replications of the consummatory response to the reinforcing stimulus. By the consummatory response is meant the complete pattern of activity commonly associated with the appearance of the reinforcing stimulus. A learned response which is a replication of the consummatory response is defined, ostensively, as one which can be measured on the same dimensions as the consummatory response. The conditioned salivary response is taken as a replication of the unconditioned salivary response to food.

[1] The study reported in this paper was carried out during the writer's tenure as research professor in the Graduate School of the University of Missouri. The research was supported in part by a grant from the University Research Council of the University of Missouri. The writer is now at the Human Resources Research Center, Lackland AFB, Texas.

In the study to be reported here we have been concerned with a learned response which is not a replicate response. The response, however, is one where increased strength of a replicate response should make for more rapid acquisition of the non-replicate response. The learning task which has been chosen is that involving the choice response in a simple T-maze. We shall consider now those implications of the R hypothesis which, taken together with certain other theoretical and empirical principles, indicate that with different amounts of reinforcement there will be observed different amounts of learning in a simple T-maze.

When in a given situation an organism responds in a certain fashion, this response becomes learned with respect to the situation if certain factors are operating. There is no general agreement among learning theorists as to all the factors which are relevant. As a matter of fact, because of the fragmentary and schematic character of a great deal of learning theory any writer hesitates to state that a given theoretical position ascribes relevance to any particular factor in an experimental situation. In spite of this it will be assumed throughout this discussion that most theorists are in agreement as to the two principles which are used in conjunction with the R hypothesis to generate the implications which have been tested experimentally in the study being reported.

The first of these principles is as follows: any two events in the learning situation become conjoined with increasing strength, in the main, as the

time between them decreases. The second principle states that the speed of locomotion of an animal towards a place where food is ordinarily found is, in the main, a positive function of the amount of food. These principles have been adequately demonstrated experimentally (1, 2, 3, 5, 8, 9).

The running response to food is here held to be a replication of the final response to food. The strength of the running response as measured by speed of running is, under the R hypothesis, a positive function of the amount of reinforcement. As the speed of locomotion increases the time between the occurrence of the turning response and reinforcement decreases.

The speed of acquisition of the turning response is a positive function of the delay of reinforcement. The speed of acquisition of the turning response will be a function then of the amount of reinforcement.

Purpose

The main purpose of the experiment has already been stated above. It may be restated here in terms of certain specific questions. These follow.

1. Will a group of albino rats rewarded with a larger amount of food learn a simple T-maze more rapidly than a group rewarded with a smaller amount of food?

2. Will the larger reward group show faster running times in terms of (1) total running time; and (2) running time from choice to endbox?

Procedure

Apparatus.—A simple T-maze was employed. This was placed on a long table. A four by eight foot copper screened panel was attached to the table with the starting alley of the T led through a small opening in the screen. Wooden uprights were nailed to the ends and sides of the T and black cloth, one yard in width, attached to these and the sides of the screen. Both arms of the T and six in. of the starting alley were within the area so screened. The maze was lighted by two double tube fluorescent lights. When the maze was illuminated the copper screened panel functioned as a one way screen.

The starting alley was three feet long. The arms of the T, measured from the sides of the starting alley, were four feet long. A non-retrace door was located at the end of the starting alley, at the entrance to the main body of the T. Two non-retrace doors were placed in the arms of the T, one foot from the choice area as measured from the sides of the starting alley. The doors were of the guillotine type. The maze was constructed of clear, finished, white pine, unpainted. The floor and walls were made of one by six in. stock, with the outer walls flush with the sides of the floor. The non-retrace doors were made of unpainted plywood.

At the end of the alleys were wooden food receptacles. These were made of blocks of finished two by two inch stock, unpainted. A circular recess, one and a half inches in diameter and a half inch in depth, was drilled in each block. The blocks were placed on the floor of the alley. It seemed unnecessary to shield them. There were no curtains in the maze.

Training.—The animals were handled before training enough to reduce wildness. They were given no pre-training.

First day: On the first training day the animal was placed in the starting box. All doors were raised. A smear of wet mash was placed on each food receptacle. The animal was allowed to move along the starting alley and into the choice area. When the animal moved through either door at the choicepoint the door was closed behind him. The animal was left in the maze until he had moved to the end of the alley and eaten the wet mash.

Ten min. later the animal was given a second trial. On this trial the animal was forced to the side opposite to that chosen on the first trial. A smear of wet mash was placed on the food receptacle and a dry pellet of food in the receptacle. When the animal had eaten he was removed. This procedure was repeated a second time 10 min. later.

Following the first three trials the animal was given a fourth trial. This was a forced run to the side chosen on the first trial. There was no food in the receptacle.

Second day: On this and all succeeding days training was by the non-correction method with equation of reinforced and non-reinforced runs. On the first two daily trials the animal was permitted to run either to the right or left. On the two following trials the animals were forced either to the right or the left. The animal was forced so as to equate reinforced and non-reinforced trials for each training day. The forced

trials were ordered so as to avoid a RLRL or LRLR sequence for the day's four trials. Spacing between trials was approximately 10 min. A food pellet was placed on one side of the maze. For any given animal this side was that opposite to that chosen on the first trial on the first day.

Maintenance and reward.—All animals were fed 12 gm. daily. This included the amount fed during training. The animals received the daily ration following training.

Prior to training the animals were assigned at random to two groups. During training one group was reinforced with a 30-mgm. pellet, the other with a 160-mgm. pellet. Training was continued for 20 days.

Recording of data: As noted above the animal was judged to have chosen one side or the other of the T when he had moved through the door on the chosen side. The door was located 12 in. from the choicepoint as measured from the side of the starting alley.

Two recordings of time were made. The first was of the interval between the placement of the animal in the maze and the animal's arrival at the end of the T following an incorrect choice, or the animal's taking the pellet from the food receptacle following a correct choice. This will be termed the total running time. It was recorded on a Standard Electric Co. timer.

The second interval recorded was that between the closing of the non-retrace door on the side chosen by the animal and the completion of the trial. This has been taken as a relative measure of delay of reinforcement. It is somewhat less than the delay period between the choice response and the appearance of the reinforcing stimulus. This interval was timed with a stopwatch.

Subjects.—The subjects were 28 male albino rats 150–200 days old, from the animal colony maintained by the department of psychology of the University of Missouri.

Results

Percents correct for free choices for all animals are given in Table I. The table has three classifications. The first is for free choice trials beginning with the 5th trial and ending with the 34th. This is a total of 16 free choice trials. The second is for free choice

TABLE I

Percents Correct for Large and Small Pellet Group Animals for Trials 5–34, 37–78, and 5–78 together with Means and Variances and *t* Ratios for Differences between Group Means for the Three Separate Trial Groupings

Groups	Trials					
	5–34		37–78		5–78	
Percent Correct	Small	Large	Small	Large	Small	Large
100			3	6		
95–99			1	4		
90–94			2	4		
85–89			1			2
80–84	1	1	1		3	7
75–79		1	1		4	4
70–74			2		1	1
65–69		4	1		2	
60–64	2	1			2	
55–59	4	4	1		1	
50–54	7	1				
45–49						
40–44		2			1	
Mean	56.25	61.19	81.22	96.19	70.64	81.35
σ^2	41.32	126.94	301.43	15.41	128.64	18.61
t	.912		2.223*		2.334*	

* Significant above the 5% and below the 2% levels of confidence.

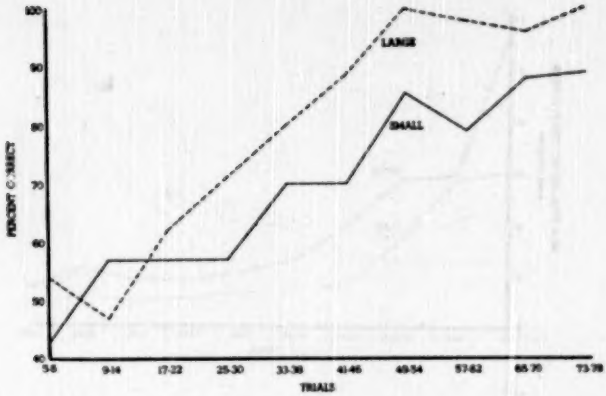


FIG. 1. Percent correct for free choice trials for the large- and small-pellet groups.

trials beginning with the 37th trial and ending with the 78th. This is a total of 22 free choice trials. The final classification is for all free choice trials. Means and variances are given together with *t*-ratios for the differences between means for the large and small pellet groups. Differences between group means for all trials and for trials 37 through 78 are significant above the five and below the two percent levels of confidence.

In Fig. 1 percents correct for free choice trials have been plotted by intervals of four free choice trials, save for the first interval. This latter interval includes the fifth and sixth trials. The curve for the large pellet group covers that for the small pellet group at every interval from the third through the tenth.

Fig. 2 shows three total running time curves. These are for all free choice trials, (A), free choice trials which were reinforced, (B), and forced choice trials which were reinforced, (C). In all of these the curves for the large pellet group are below those for the small pellet group, except for the second to third trial interval for the reinforced choice trials.

Fig. 3 shows choice to endbox running time for all free choice trials, (A), for reinforced free choice trials, (B), and for reinforced forced choice trials, (C). Curves for the large pellet group are below those for the small pellet group at all intervals save that for the second interval for reinforced free choice trials.

No attempt has been made to establish the reliability of any difference between mean running time for groups. No inferences as to population parameters have been made.

DISCUSSION

The most notable feature of the results just described is the presence of a difference in learning between the two experimental groups. This difference is in the direction predicted on the basis of the R hypothesis.

Differences in learning measured in terms of frequency are paralleled by consistent differences in running time between groups. These differences in running time are all in the same direction as the differences in frequency. Of particular interest are the running times following the choice responses. All running time curves

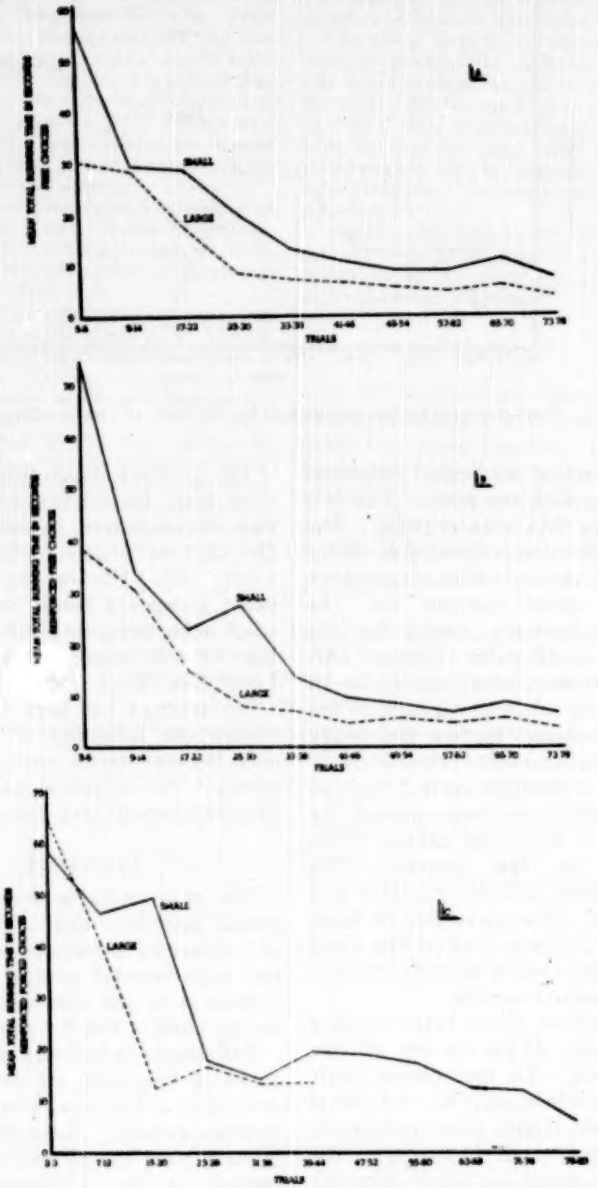


FIG. 2. Mean total running time curves for large- and small-pellet group animals for (A) all free-choice trials; (B) free-choice reinforced trials; (C) forced-choice reinforced trials

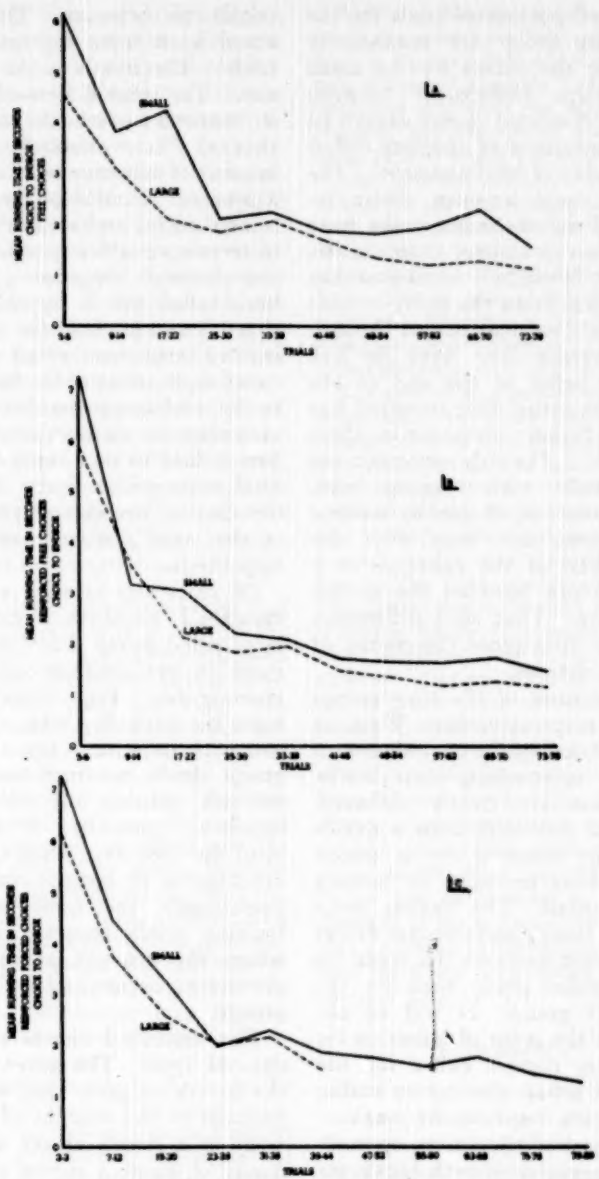


FIG. 3. Mean choice to endbox running time for (A) all free-choice trials; (B) reinforced free-choice trials; and (C) reinforced forced-choice trials

for the final portion of trials for the large pellet group are consistently covered by the curves for the small pellet group. Differences between these sets of related curves cannot be taken as measures of absolute differences in delay of reinforcement. The measured times between choice response and end of the run in the maze are consistently shorter than the delay times. It will be remembered that recording was from the point in time when the animal had passed through the non-retrace door until he had taken the pellet at the end of the maze. The actual time recorded was that for a distance about six in. short of three feet. The only concern in the present study with running time, whether total or choice to endbox running time, has been with the determination of the existence of a clear difference between the groups under study. That such differences are present throughout the course of training is obvious.

Three features of the time curves are worth some attention. First, at the end of training the curves seem to be closely approaching their limits. These limits are clearly different. Second, the curves all show a sizable initial drop followed by a period where absolute decreases in running time are small. The flexion point separating these phases occurs earlier and at a point lower on the *y*-axis for the large pellet group than for the small pellet group. It will be observed that the point of inflection for the percents correct curve for the large pellet group also occurs earlier and at a point lower on the *y*-axis.

Third, most of the curves approximate the negative growth curve for the general function $t = Pe^{-rN}$, where t is running time, N is trials, and e is the base of natural logarithms. P is taken in terms of some measure of initial performance. This will be based upon some aggregate of early trials. The term r is the rate of decay. The general form of the curves is consistent with the assumptions that (1) P is a variable function of the amount of reinforcement and (2) r is a constant function of the amount of reinforcement and a variable function of certain variables specific to a given experimental situation. What has been called the R hypothesis states that (1) and (2) hold for that class of learned responses which have been called replications of the final response to the reinforcing stimulus. The running response we have been discussing here is held to be a replication of the final response to food. The results obtained in the case of running time in the maze positively verify the R hypothesis.

Of particular interest are the unsmoothed acquisition curves. The large pellet group shows no improvement in performance until the 5th training day. Improvement is rapid until the 13th day when all animals run without error. The small pellet group shows no improvement until the 9th training day. Performance improves gradually, if irregularly, until the 16th day, when the animals are running 89 percent correct. Approximately the same number of training trials separate the points where the greatest amount of improvement begins and ends for both groups.

The smoothed curves are of the sigmoid type. The general form of the curves suggests that with manipulation of the amount of reinforcement in a simple choice situation a family of sigmoid curves will be obtained, with a negatively accelerated curve as the limiting form with increasing amounts of reinforcement and a straight line with zero slope as

the limiting form with decreasing amounts of reinforcement. Presumably with increased amounts of reinforcement the limit for the percents correct curve would reach 100 percent. No further increase would then be possible. If we assume that all curves would conform to the general equation $P = 50(100\text{-}b)r^N$, where P is percent correct, b a parameter, r the rate of growth, and N, trials, then (1) r would be a function of the amount of reinforcement and (2) b would be a function of the amount of reinforcement within the limits 0–50.

In considering the percents correct curves the question arises as to whether it is possible to ascribe any meaning to the flexion points of the curves. This seems possible in terms of the R hypothesis. This interpretation must be tentative only.

The form of the percents curves is that commonly observed with many cumulative curves which represent the measured effects of the interaction of individual differences with other variables. The variables interacting with individual differences in the present instance may be considered to be those related to the replicate response. As the replicate response moves back in the response sequence it becomes integrated with the choice response. The choice response becomes a component of the total response pattern to food, differentiated with respect to a particular situation. The rate of backward movement of the replicate response is a function of the amount of reinforcement (1). Integration will occur earlier in time as reinforcement increases. What is here termed integration may be given empirical meaning in terms of the point of inflection of the percents correct curve. The point of inflection represents the beginning of the integration of the choice response into the goal response for the group under training. The limit of the curve represents the completion of the integration process for all animals.

Wolfe and Kaplon (9) have reported a study the results of which are in agreement with those reported here. Chicks were trained in a simple T-maze. Reinforcement was with one whole grain of popcorn, one one-fourth grain of popcorn, or one whole grain cut into four pieces. Fastest running time was observed for the four-piece group, slowest with the quarter-grain group. Fewest errors were made by the four-piece group, most by the quarter-grain group. No results of tests of significance were reported. Extreme variability in performance was observed. Presumably error differences were not significant.

The results of a second study in the same report by Wolfe and Kaplon are consistent with the R hypothesis. In this study a problem box with a glass partition was used. The chicks were required to learn to go around the partition to food. The four-piece group was superior to the other groups in running time in the final experimental sessions. The writers report that this result, when it became known to them after the study was completed, was a source of some surprise. During training more nonadaptive responses had been observed among four-piece group animals than among the other animals. The animals of the four-piece group stayed nearer the food, pecked at the glass longer, and showed a more restricted variety of responses during the presolution period than did the other animals. This behavior can be ascribed to the greater strength of the replicate response which was evidenced in the tendency to maintain the direct approaching response to the goal object.

Summary

1. Two groups of 14 male albino rats were trained in a simple T-maze. One group was rewarded with a 30-mgm. pellet of food, the other with a 160-mgm. pellet. Four trials, two free-choice and two forced-choice runs, were given each training day. Training was continued through 20 days.

2. In terms of percents correct, the large-pellet group showed a greater amount of learning. The percents correct cuve for the large-pellet group showed a faster rate of growth and a higher limit.

3. The large-pellet group showed faster running times for free-choice trials and for reinforced trials.

4. The above results were held to be consistent with what has been termed the R hypothesis.

(Manuscript received April 11, 1949)

References

1. Crespi, L. P. Quantitative variation of incentive and performance in the white rat. *Amer. J. Psychol.*, 1942, 55, 467–517.
2. Geier, F. M., & Tolman, E. C. Goal distance and restless activity: I. The goal gradient of restless activity. *J. comp. Psychol.*, 1943, 35, 197–204.
3. Hull, C. L. The rat's speed-of-locomotion gradient in the approach to food. *J. comp. Psychol.*, 1934, 17, 393–422.
4. Perin, C. T.: The effect of delayed reinforcement upon the differentiation of bar responses in white rat. *J. exp. Psychol.*, 1943, 32, 95–109.
5. Reynolds, B. The acquisition of a trace conditioned response as a function of the magnitude of the stimulus trace. *J. exp. Psychol.*, 1945, 35, 15–30.
6. Reynolds, B. The acquisition of a black-white discrimination habit under two levels of reinforcement. *J. exp. Psychol.*, 1949, 39, 760–769.
7. Reynolds, B. Resistance to extinction as a function of the amount of reinforcement present during acquisition. *J. exp. Psychol.*, 1950, 40, 46–52.
8. Wolfe, J. B. The effect of delayed reward upon learning in the white rat. *J. comp. Psychol.*, 1934, 17, 1–21.
9. Wolfe, J. B., & Kaplon, M. D. Effect of amount of reward and consummative activity on learning in chickens. *J. comp. Psychol.*, 1941, 31, 353–361.

DRIVE SPECIFICITY AND LEARNING: THE ACQUISITION OF A SPATIAL RESPONSE TO FOOD UNDER CONDITIONS OF WATER DEPRIVATION AND FOOD SATIATION

BY EDWARD L. WALKER, MARGARET C. KNOTTER, AND RUSSELL L. DEVALOIS

*University of Michigan**

A large number of experiments have appeared in the literature in recent years which have concerned themselves with such topics as 'drive discrimination,' 'drive interaction,' 'drive specificity,' 'motivational control of habits,' 'latent learning,' 'sign Gestalt expectations,' 'secondary reinforcement,' and a multitude of others. On a theoretical level these studies have been designed to investigate many topics and problems. On an empirical level a large proportion of them are related to a single problem which might be stated as follows: *To what degree must an incentive be relevant to a drive for learning to occur?*

In its fullest elaboration this is an extremely complex problem involving the determination of the degree of 'relevance' of a vast number of possible incentives and possible drives as well as the effect of a large number of other variables on the phenomenon. Actually we do not as yet have adequate answers to the simplest form or condition of the above question: that is, the drive-incentive relationship necessary for learning when most of us can agree that a given incentive is relevant to a particular drive and not to another. If we disregard the complexities, we can elaborate the question in this form: *Can an accepted or standard incentive, reward or goal object be demonstrated to play an effective role in the learning situation when the operations which are employed to motivate the animal are: (1) clearly relevant to the incentive, (2) clearly irrelevant, or (3) when there is no drive state, either relevant or irrelevant?*

That the answer to this question is 'Yes' when the drive is relevant to the incentive cannot be questioned in view of the great number of experiments involving food deprivation and food reward, for example.

Whether learning can take place when there is no drive is almost a meaningless question since there is no clear set of operations for producing a state of 'no drive.' It was included here partly for the sake of logical completeness and partly because attempts to create this condition have led to data pertinent to the second alternative.

The answer to this question in the situation in which it is clear that the drive is irrelevant to the incentive is complex. In an experiment by Spence and Lippitt (3), which has been successfully repeated a number of times, animals were satiated for food and water and were run in a single choice point maze with one choice leading to food and the other to water. After a period of training they were made hungry or thirsty and demonstrated appropriate responses —the hungry animals went to food and the thirsty animals went to water. These experimenters found it necessary to motivate their animals in some way during the initial training to get them to run at all, and they rewarded running behavior by placing

* The work herein reported was supported in part by a research grant from the Horace Rackham School of Graduate Studies at the University of Michigan.

the animal in a social cage at the end of the run. Thus the learning that was demonstrated in the hunger and thirst motivated trials was acquired under a *weak*, irrelevant drive—a social, gregarious, or isolation-avoidance drive.

Unfortunately, there are at least four studies reported in the literature which seem to indicate that such is not the case when the irrelevant drive is *strong*. Spence and Lippitt (4) and Walker (5) trained animals under thirst motivation and rewarded one turn with water and permitted animals to experience food (the irrelevant incentive) after the other turn. When the animals were made hungry, they continued to go to water and seemed not to have learned the location of or acquired a response tendency to food.

It occurred to many people that the strong water-going response set up in these experiments might have prevented the demonstration of a weaker food-going tendency. Kendler and Mencher (2) and Grice (1) therefore rewarded either turn with water during training when the animals were thirsty and placed food on one side only. When these animals were subsequently made hungry, they failed to exhibit a statistically significant tendency to go to food.

The present experiment was set up to test the same hypothesis: *Rats can learn the location of food or acquire a response tendency associated with food when training is carried out under strong thirst motivation and when at the time of test under hunger motivation there is no other strong response tendency to either of the alternative turns.*

Since in our present experimental design there was one source of positive response tendency, position preference, which was controlled but not eliminated, we had an additional hypothesis; namely, *the weaker the position preference, as measured, the more likely the demonstration of a food-going response tendency*, since any food-going response established in the training would have to compete with any natural or existing position preference or turning preference that might characterize that particular animal.

Both of these hypotheses were confirmed in the present study, which was performed in part before the appearance of the Kendler and Mencher and Grice papers. Since their results were negative and ours positive and since they used 50 percent forced trials during training and we did not, two additional repetitions of our original experiment were performed and are here reported using 50 percent forced trials in one instance and 100 percent forced trials in the other.

EXPERIMENTAL PROCEDURE

Subjects.—A total of 66 naive rats were used in these experiments. They ranged from 100 to 138 days of age at the beginning of training. The 26 animals in the first experiment were half male and half female and were drawn from the Maier-Shepard colony at the University of Michigan. They were almost exclusively albino. The 40 animals used in the second and third experiments were all male Wistar albinos.

Apparatus.—A diagram of the apparatus is shown in Fig. 1. The dimensions in inches are as follows: width of alley—5; height of alley—5; length of starting box—9½; length of starting alley—36; sides of choice point—12; top of choice point—16; length of choice alley—21; length of each goal box—15; width of goal box—15.

The maze rested on a sheet of linoleum painted gray. It was constructed of aluminum sheet metal with a hardware cloth top throughout. The bottoms of the choice point and goal boxes were metal, all other sections being open. Near each of the first six of the eight string-controlled guillotine type gates (indicated by triple lines in the diagram) there was a piece of ¼-in. brass rod, ¼ in. from the floor, inserted to brace the maze at those points. The tops of the starting chamber, choice point and goal boxes were hinged. Since the floors of the choice

alleys were open, it was possible to manipulate their nature. During this experiment, there was a piece of white cardboard under a piece of fine hardware cloth under one choice alley and a piece of black cardboard under a piece of coarse hardware cloth under the other. These insertions extended from the metal floor of the choice

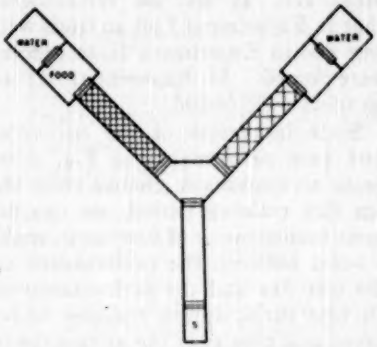


FIG. 1. Diagram of experimental maze. (See text for dimensions and detailed description.)

point to the metal floor of the goal boxes, thus from gate to gate. In addition the lights over the maze were so arranged that there was a wide but not dark shadow in the alley with the white cardboard and a narrow but not dark shadow in the alley with the dark cardboard.

Familiarization and test for position preference.—Day 1—With neither water nor food in the maze and with the animals relatively satiated for both, they were introduced into the maze one at a time and the gates closed behind them after they had made a choice. They were then removed from beyond the choice point and reintroduced later at the point of removal with several other rats. Each group was allowed to remain in the maze with all gates open for a familiarization period of 20 min.

Days 2 and 3—Each animal was given two trials each day. The two trials on a given day were separated from each other by approximately 90 min., the amount of time required to run all other animals being trained at the same time. Drive-incentive conditions were the same as on Day 1.

Learning series.—Animals were given four trials per day under 16–18 hours of water deprivation and were rewarded regardless of choice in the terminal goal box. Food intervened on one side only. The water was in the usual cage bottle which was inserted through the roof of the terminal goal box. The food was a fairly large quantity of their regular laboratory chow pellets in the intervening box. Both food and water were arranged in the box so that they could not be seen from the choice point, but were quite visible when the rat was in the goal box. When the rat was placed in the maze, all gates were open, and the gates were closed behind him as he passed through each point including the gate between the goal boxes.

The side on which the food was placed was determined by the original position preference trials. For half of the animals the food was placed on the side of the animal's preference and for the other half it was placed on the side opposite from the animal's preference. In so far as strength of position preference could be determined from the five preliminary trials, an effort was made to match sides in this respect also, with every rat with a strong right preference with food on the right matched with an animal with an equally strong right preference with food on the left. In the first experiment the food was on the left for 14 animals and on the right for 12. In the other two experiments the food was on the left for half and on the right for half.

To insure satiation for food during the training trials under thirst motivation, three steps were taken. Food was present in the home cage at all times. Each animal was placed in a small cage containing a large quantity of food just before the experimental run. Lastly, each animal was fed by hand before introduction into the maze. Under these conditions no animal consumed food in the maze.

The only difference in the training procedure in the three experiments was that in the first experiment all 40 trials of training (4 trials a day for 10 days) were free, in the second experiment 50 percent (the second and fourth) were forced in the opposite direction from the preceding trial, and in the thirst experiment 100 percent of the trials were forced in random order within a given day. Forcing was accomplished by having one of the choice point gates down when the animal was placed in the maze.

Test trial.—At the end of the tenth day of training, water was restored to the home cages, all food was removed, and on the test day, each animal was given a single trial under 22 hours of food deprivation.

RESULTS

The major hypothesis under consideration in this study is: *Rats can learn the location of food or acquire a response tendency associated with food when training is carried out under strong thirst motivation and when at the time of test under hunger motivation*

there is no other strong response tendency to either of the alternative turns. The answer is essentially in the affirmative. Fig. 2 indicates that after a period of training under thirst motivation, when the motivation is changed to hunger, there is an abrupt positive change in the number of animals taking the turn leading to food.

The significance of this shift could be tested in several ways. The most severe test would be to compare the proportion of animals taking the turn to the food side on the first trial of the tenth day of training with the proportion taking the turn to the food side on the test day. These shifts were from 42 to 69 percent in Experiment I, from 35 to 65 percent in Experiment II, and from 50 to 60 percent in Experiment III. These differences as proportions and the results of the tests of significance may be seen in Table I.

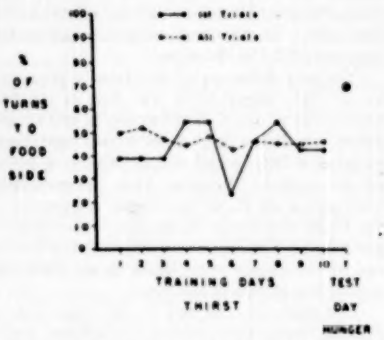


FIG. 2. Experiment I. The graph presents curves showing the performance of the 26 animals in the first experiment. The broken line indicates the percentage of all turns to the side of the maze where there was both food and water on each of the ten days of training under thirst motivation. Each point represents 104 turns (4 turns × 26 animals). The solid line represents the percentage of animals which went to the food side on the first trial only of each training day—26 turns. The double circle represents the percentage of animals going to the food side on the single trial of the test day when they were hunger motivated.

The obtained *t*-values in Experiments I and II indicate that the differences are significant at about the five percent level in each instance. The difference was not significant but in the expected direction in Experiment III. It will be remembered that in Experiment I all 40 trials were free and in Experiment II 50 percent were forced. In Experiment III all 40 trials were forced.

Since inspection of the curves of first turn performance in Fig. 2 reveals no systematic change over the ten day training period, we can assume homogeneity of first turns, make a *t*-test between the performance on the test day and the performance on all first turns during training under thirst and thus gain the advantage of an increased N and an increase in degrees of freedom. This t indicates significance at the one percent level of confidence in Experiment I. Various other tests can be made, this same test in Experiment II, test performance as against total free trials, etc., but they lead only to additional *t*-values falling generally between the five and one percent levels of confidence. On these bases it seems reasonable to reject the null form of our major hypothesis.

The percentage of animals taking the turn to food on the test trial is decreased as the number of free trials is decreased; that is, from 69 percent with all free trials to 65 percent with half free trials to 60 percent when none are free.

It will be recalled that in the design of this experiment one major source of preferential response tendency (position preference) to one or the other side not due to differential reinforcement was controlled but not eliminated. Two controls of this natural response tendency were introduced. As noted in the procedure,

TABLE I

COMPARISON OF THE PERFORMANCE ON THE TEST DAY UNDER HUNGER MOTIVATION WITH PERFORMANCE DURING TRAINING UNDER THIRST MOTIVATION

In Experiment I all training trials were free, in Experiment II, 50% were forced, and in Experiment III all training trials were forced.

	N	Proportion of Turns to Food	Diff.	$\sigma_{Diff.}$*	*t*
Experiment I					
Test Day (Hunger)	26	.69			
First Trial Tenth Day (Thirst)	26	.42	.27	.137	1.97
Total First Trials (Thirst)	260	.43	.26	.102	2.58
Experiment II					
Test Day (Hunger)	20	.65			
First Trial Tenth Day (Thirst)	20	.35	.30	.155	1.94
Experiment III					
Test Day (Hunger)	20	.60			
First Trial Tenth Day (Thirst)	20	.50	.10	.173	.58

* This estimate of the standard error of the difference is based upon the combined proportion assuming the null hypothesis. The degrees of freedom are $(N + N - 2)$. No correlation term was used in estimating the standard error of the difference because of a question of statistical logic. The omission leads, of course, to a slight underestimation of *t* in each instance.

each animal had five operationally unmotivated and unrewarded trials in the maze before training began and in all three experiments the food was placed on the preferred side for half of the animals and on the non-preferred side for the other half. While this tended to load the two choices equally in this respect, it provided no very reliable criterion of the strength of such a tendency. Since both turns were reinforced during training under thirst, any preferential performance during this period should be an adequate measure of position preference. Further, since only the first trial in a day's training was relatively free from the well known tendency for rats to alternate in such a situation, the best available measure of strength of position preference was the performance on the first turns of the ten days of training. Thus, an animal that turned right when first placed in the maze on each of the 10 days of training would be said to have a strong right turning preference, and one which turned right nine times out of ten such opportunities would have a somewhat less strong position preference, while the animal which took five such turns to the right and five to the left would have no demonstrable preference for either side.

On the basis of this reasoning and the performance of the animals in Experiments I and II (there were no free trials in Experiment III) the data could be sectioned on the basis of the strength of position preference and our second hypothesis that *the weaker the position preference, as measured, the more likely the demonstration of a food-going response tendency* could be tested.

When the data in Experiments I and II were examined it was found that 19 of the 46 animals took as many as nine out of ten first trials during thirst training to one side or the other. These have been designated as the 'Strong' position preference animals in Table II. An additional group of 16 animals were found

to have taken either seven or eight first turns to one side during thirst training and these have been designated as animals with a 'Medium' position preference. The remaining 11 animals took either half of their first trials to each side or six turns to one side and four to the other. These have been designated as 'Weak' position preference animals. The critical statistic with reference to our hypothesis above is the percentage of each group that modified their performance when the deprivation was changed from water to food. As can be seen in Table II, almost no shift in performance is noted in the 'Strong' position preference group while a maximum shift is seen in the group with the 'Weak' position preference.

TABLE II

AMOUNT OF CHANGE IN PERFORMANCE WITH CHANGE IN MOTIVATION RELATED TO STRENGTH OF POSITION PREFERENCE

See text for criterion of strength of position preference. These data are taken from Experiments I and II only since no measure of position preference is possible in Experiment III.

Degree of Position Preference	*N*	Percent of Animals Modifying Performance
Strong	19	5.3
Medium	16	37.5
Weak	11	54.5

Two additional facts might be noted in interpreting these results. The side preferred appears from the data to be related primarily to the absolute position (right or left) more than to the presence of food which was on each side for roughly half of the animals. Second, the two measures of position preference used here are virtually uncorrelated.

DISCUSSION

The original empirical question with which we started was: *To what degree must an incentive be relevant to a drive for learning to occur?*

If we can accept food as an incentive which is clearly irrelevant to the drive produced by water deprivation, then the results of this experiment in conjunction with those of Spence and Lippitt (3) seem to admit of the interpretation that learning can occur when the drive and the incentive are clearly irrelevant.

This interpretation might be challenged on a variety of grounds. First, when part of these results was reported at the annual meetings of the American Psychological Association in Boston in September, 1948, Kendler suggested from the floor that the results might be accounted for in another way. He suggested that when the drive state was changed many animals might simply have done something else in the choice situation other than what they had last done in that situation. Since at the end of training under thirst the animals in Experiment I were taking only 42 percent of their first turns to the food side, such a shift would lead to an expectation of 58 percent turns to the food side without food having played any role in the training situation. This interpretation must be rejected, since of the 46 animals in the first two experiments, 33 made the same turn under both drive states and 13 shifted from the non-food to the food side. Not a single animal shifted from the food to the non-food side.

Second, there are the negative results in the Kendler and Mencher (2) and the Grice (1) studies. Their procedures were quite similar to the present Experiment II in which the results were positive. There is no

really clear basis for a guess as to why their results were negative. There is a possibility that had their training periods been extended or had the choice been more discriminable, they would also have obtained evidence of an acquired food-going response tendency. Regardless of such speculation, it must be remembered that within the logic of statistical tests, a failure to reject the null hypothesis does not in any sense imply acceptance of it, and when all three results are taken together, we must conclude that food can serve as an effective incentive or secondary reinforcer under thirst.

A third class of challenge involves the relevance of food to the thirst drive. The food in this instance consisted of dry laboratory pellets and consumption of them when thirsty, an event which did not occur, certinaly would not have led to a reduction of the primary thirst drive, but more likely to an increase in it. In this sense we can probably all agree on the question of relevance. In a more subtle sense the question might be raised that the sight of food was often in the life history of the animal present at the time of alleviation or reduction of the thirst drive and this acquired secondary reinforcing properties *in relation to the thirst drive.* Such an interpretation would require that the sight of food be largely limited to the period of thirst reduction; otherwise the acquired secondary reinforcing properties would be extinguished. The plausibility of this interpretation is not great and is even less so in the Spence and Lippitt situation where the irrelevant drive was a social drive and the presence of other animals is virtually continuous in the life history of the animals.

Last, there is the possiblity that our procedures did not produce absolute satiation of the hunger drive and that the observed food-going response tendency was acquired under a weak hunger drive which is ineradicable under strong thrist. For the moment, at least, we must say that the hunger drive was operationally absent as a result of these satiation procedures, and secondary or supportive evidence is available in that the animals did not eat in the maze, even after as many as three water rewards.

Having rejected all four classes of objection, we feel that the original interpretation of learning when the drive and incentive are not relevant to each other stands.

The secondary hypothesis concerning the interaction of the acquired food-going tendency and the position preference of the animal simply gives second-hand evidence of the relative strength of a response tendency acquired under an irrelevant drive as compared to one acquired under a relevant drive. The period of training used in this study is adequate for the acquisition of a water-going response tendency when thirsty animals are rewarded with water (4, 5) which is strong enough to overcome any strength of position preference demonstrated by the animals used in those experiments. The food-going response tendency demonstrated in the present experiments is not as strong as the strongest position preference.

The role of forced trials is an interesting problem for which this series of experiments offers no definitive answer. Forcing all trials during training seems to have reduced the strength of the acquired food-going response tendency, and in a statistical sense we are not justified in saying that any such response tendency was acquired in Experiment III. Whether forced trials affect what is learned directly, whether they operate to change the conditions of motivation, or whether

they produce something like frustration is an area yet to be investigated systematically.

To sum up, it seems to the present author that, in terms that Hull might use, the sight of food has served here as a secondary reinforcement and increments of ${}_{s}H_{R}$ to the food side have been acquired during training under thirst motivation. Such acquired habit strength remained unexpressed in effective reaction potential until the drive state and thus the drive stimuli of hunger were present.

In terms that Tolman might use, we have the acquisition of a sign-Gestalt expectation of food in such-and-such a place independent of 'reinforcement' and a clear demonstration of the distinction between learning and performance as this distinction is usually made by Tolman and his followers. In short we have 'latent learning.'

Perhaps it is possible to accept both statements.

SUMMARY

A total of 66 young, naive rats were given four trials of training per day for 10 days under 16–18 hours of water deprivation and food satiation in a single choice point Y-maze. They were rewarded with water on both sides and food was placed on one arm of the maze only. For 26 animals, all 40 training trials were free, for 20 animals 20 trials were free and 20 forced, and for 20 animals, all 40 trials were forced. On the 11th day each animal was given one test trial under 22 hours of food deprivation.

A significant tendency for the animals to go to food when made hungry was demonstrated in both groups with free trials and a similar but not significant tendency was demonstrated in the group in which all trials were forced.

A secondary hypothesis that the weaker the position preference of the animal, the more likely the demonstration of an acquired food-going response tendency was also confirmed.

It is suggested that the results might be acceptably expressed in the terms of either Hull or Tolman.

(Manuscript received April 4, 1949)

REFERENCES

1. GRICE, R. G. An experimental test of the expectation theory of learning. *J. comp. physiol. Psychol.*, 1948, 41, 137–143.
2. KENDLER, H. H., & MENCHER, HELEN C. The ability of rats to learn the location of food when motivated by thirst—an experimental reply to Leeper. *J. exp. Psychol.*, 1948, 38, 82–88.
3. SPENCE, K. W., & LIPPITT, R. O. "Latent" learning of a simple maze problem with relevant needs satiated. *Psychol. Bull.*, 1940, 37, 429.
4. SPENCE, K. W., & LIPPITT, R. O. An experimental test of the sign-gestalt theory of trial and error learning. *J. exp. Psychol.*, 1946, 36, 491–502.
5. WALKER, E. L. Drive specificity and learning. *J. exp. Psychol.*, 1948, 38, 39–49.

STUDIES IN MOTIVATION AND LEARNING: II. THIRSTY RATS TRAINED IN A MAZE WITH FOOD BUT NO WATER; THEN RUN HUNGRY[1]

BY HENRY GLEITMAN

From the Psychological Laboratory, University of California

Introduction

The sign-Gestalt or expectancy theory of learning (6) has been challenged recently by a number of studies carried out by Spence and Lippitt (4), Kendler (1), Kendler and Mencher (2), and others, who failed to obtain latent learning in one particular type of situation. Generally speaking, they found that thirsty rats run in a simple T- or Y-maze containing water on one side and food on the other did not learn the location of food during this training situation, as evidenced by their failure to run to the food side when subsequently deprived of food. In this type of situation no latent learning was demonstrated. The results of this series of experiments were interpreted by Spence and Lippitt (4) and by Kendler (1) as presumptive evidence against the existence of latent learning in general, and were considered as favorable evidence for reinforcement theory.

A recent experiment, performed by Tolman and Gleitman (8), however, has shed some doubt on such an interpretation of these findings. A particular maze situation was found, in which the animals, though being rewarded equally on both sides, were still able to learn the location of two very strongly differentiated goal-boxes, as demonstrated by their behavior when placed on the maze after having been shocked in one of the end-boxes in another room. These results suggested that while latent learning may not occur under all conditions, particularly not those utilized in the experiments performed by, or modeled after, that of Spence and Lippitt, there are some others under which it will. Nevertheless it still remains for the expectancy theory to embark upon a program of theory construction which will permit the more specific prediction of the particular conditions that permit or prohibit the occurrence of latent learning.

Certain hypotheses may tentatively be advanced as to the nature of the factors that influence such learning. Thus it is believed that perceptual 'outstandingness' of features in the field, and the degree of motivation (during the original perception of the situation, during later performance, and possibly during the retention period) may all be of importance in determining the type of 'cognitive map' (7) that the animal will form and later be able to act upon. At present, however, the assumption of the importance of such factors is more in the nature of a speculation than of a well organized hypothesis, and perhaps for the time being, the course of the necessary theory construction will be better served by a program of systematic investigation of the factors involved in latent learning—both those producing it and those preventing its occurrence—than by a too elaborate formulation of further theory. One of the more specific problems which should be attacked first is a further investigation of the same situation utilized by Spence and

[1] The writer is deeply grateful to Professor Edward C. Tolman for generous advice and aid.

Lippitt with an attempted analysis of the factors which make for the absence of latent learning in this particular set-up.

The present paper is an attempt to check on two possible interpretations of the results obtained by these authors. (1) One might argue that the rats really *knew* the location of the food, but continued running to the water side, because the side containing food had led to negative emotional reactions during the original training period, possibly induced by the method of forced trials. In an unpublished repetition of the original Spence and Lippitt experiment run in this laboratory by R. Schweers, considerable escape activity and emotional behavior were observed in the alley leading to the inappropriate goal-object, when the animals were forced to run there. This would seem to lend some degree of a priori plausibility to such an interpretation. (2) A second hypothesis would start from the assumption that the animals really did *not* learn the location of the food. The reason for the absence of latent learning would be sought in the presence of the appropriate goal-object—water—on the maze during the original training situation. One might hypothesize, that is, that the absence of learning was due to the distracting effect of the strongly valenced water which would narrow the interest and therefore the resulting 'cognitive map' of the animal, so that all things not relating to this goal-object would tend to be disregarded. In short, perhaps the animal would have learned the location of the inappropriate goal-object—food—if the appropriate one —water—had not been in the field.

If either of these two hypotheses were correct, one might then expect the animals to learn the location of the food, if they were run thirsty in a maze containing food in one end-box, but neither food nor water in the other. In such a set-up, neither would any escape behavior be induced on one side and not on the other, nor would the presence of the appropriate goal-object create a narrowing of the animal's cognitive map. If negative results are obtained in such a set-up, these two hypotheses at least will have been ruled out, and the field of possible explanatory factors narrowed.

Apparatus

The apparatus used consisted of a Y-shaped alley-maze. The alleys were two feet long, six in. high, and five in. wide. Two end-boxes, 10 by 14 by 10 in., were located at the end of two of the alleys, alleys B and C. The floors of both end-boxes were covered with saw-dust. One-way hanging doors, which the rat manipulated himself, made of black cardboard, were placed, one in the approach alley, four in. from the choice point, two others five in. away from the choice point in both alleys B and C, and one each at the entrance to each end-box.

To provide a greater diversity of directing cues, the left side of the choice point area and the floor of alley B (leading to the empty goal box) was covered with coarse hardware cloth, while the right side of the choice point area and the floor of alley C (leading to the end-box containing food) was covered with fine hardware cloth.

The room was illuminated by six ceiling lights, the maze being placed symmetrically with respect to the lighting. The experiment was conducted from about 2:00 A.M. to 4:00 A.M. daily.

Procedure

1. *Pr.'iminary training.*—Twenty-four pigmented male rats from the U. C. laboratories (Tryon's M by M stock) were used for this experiment, ranging in age from four to six months. None of the animals had any previous maze experience. To accustom them to the apparatus they were permitted to live in the maze for several nights. During this preliminary training the one-way doors were not installed in the maze, nor was any food or water placed in the end-boxes. Two days after this procedure, each rat was given four free runs through the maze, to determine his position preference. During this part of the training also neither food nor water was placed in the maze. However, the one-way-doors are installed at this point, and remained in the maze for the rest of the experiment.

On the basis of the results of the four preliminary trials, the animals were divided into two groups of 12 rats each, in such a manner that an equal number of animals with right or left preferences were in each group.

2. *Training under conditions of thirst motivation.*—During the next five nights, both groups of animals were run under 22 hours of water deprivation. They received four trials a night, of which one was fee and the other three were forced, in such a manner that each animal ran twice to each side every night. During the free trials the one-way doors could be easily opened by a slight push of the animal. During the forced trials, on the other hand, one of the two doors at the choice point was held shut with a string, forcing the animal to run to the opposite side. The free trial was given first each night.

Group I (Experimental Group) always found food in the right hand goal-box and an empty goal-box on the left, while Group II (Control Group) found food in the right hand goal-box and water on the left.

Food was placed in a food bin in the right hand goal-box identical with that which the animals fed from in their living cages, and water was found through a hole in the left hand goal-box, through which the muzzle of a regular drinking bottle protruded when Group II was run. The food consisted of the regular dry powdered diet normally fed the colony.

To insure objective equality of experiences in the end-boxes the animals were detained there for 30 sec. after each run. If the animal hesitated for more than six min. at the choice point, he was removed from the maze, and the trial repeated at a later time.

Animals in both groups were run satiated for food. This was accomplished by having food available to them at all times in their living cages and by filling the food bin in their living cages to the brim about two hours before running. Upon completion of their four trials the animals were watered in their living cages.

3. *Critical test and learning under hunger motivation.*—After the fifth night the animals in each group were permitted free access to water at all times, but were placed on a food maintenance schedule. To insure a hunger rhythm, the sixth calendar night was skipped and the animals run on the seventh under 22 hours of food deprivation, but satiated for water. Their bottles in the home cages were always completely refilled about two hours before running during this phase of the experiment.

On both the seventh and the eighth night each of the animals was given six free trials per night. The content of the goal-boxes was identical with that in the previous part of the experiment—food being on the right for both groups, but water present on the left for Group II and absent for Group I.

The animals were kept in the goal-boxes for 20 sec. after each trial.

Results and Discussion

As shown in Table II, on the first critical trial under hunger motivation, all of the animals in Group II (Control Group) continued to run to the left side, the side on which they had found water during the original training situation. This once more sub-

TABLE I

Choices in Original Training
Thirst Motivation
No. of Right Hand Choices on Free Trial

Day	Group I—Exp.	Group II—Control
	$N = 12$	$N = 12$
1	6	8
2	6	0
3	5	0
4	6	1
5	2	0
	Food on right Empty on left	Food on right Water on left

TABLE II

Choices in Subsequent Training
Hunger Motivation
No. of Right Hand Choices

		Group I—Exp.	Group II—Control
		$N = 12$	$N = 12$
Day 7	Trial 1	7	0
	2	9	10
	3	12	10
	4	11	7
	5	8	11
	6	11	8
Day 8	Trial 7	12	10
	8	7	7
	9	9	9
	10	12	10
	11	8	11
	12	10	11
		Food on right Empty on left	Food on right Water on left

stantiates the results of Spence and Lippitt. It is seen, furthermore, that Group I (Experimental Group) did not shift significantly to the food side on this trial. While there is a statistically significant difference ($t = 2.02$) for this group between the number of animals choosing that side on the critical trial and the number of animals choosing it on the free trial of the fifth night—the last night under thirst motivation (see Table I and Table II)—it seems more reasonable to ascribe this difference to a fluctuation on the fifth night rather than to attribute it to latent learning, for the difference between the fifth and fourth night performance is also significant for this group ($t = 2.1$). No significant difference exists between the performance on the critical trial for Group I and that of the free trials of the fourth, third, second, or first nights of the training period.

A comparison of the learning rates under hunger motivation for both groups during the next eleven trials (see Table II) is a further indication that the animals in Group I (Experimental Group) had not learned the location of food during the original training period, any more than had those in Group II. According to the hypotheses to be tested, the presence of the original goal-object—water—should have been a disturbing influence on the acquisition of knowledge regarding the food's location—either by creating fear, or by a narrowing of the animal's perceptual awareness. We should, therefore, expect better learning scores for Group I, the group which found only an empty end-box on one side, than for Group II, the one which found water there. If Group II had learned something more, it should have showed up in such a comparison. However, a comparison of the learning scores shows that the mean number of errors per rat for all 11 trials is 2.42 for Group II and 2.00 for Group I. This difference is not significant, yielding a t-value of only 1.05. These results make it reasonable to suppose that the animals in Group I had not learned the location of the food any more than had the animals in Group II, which were run under the original Spence and Lippitt set-up.

These findings, while not permitting any positive statements regarding the specific conditions that permit or prohibit latent learning, present, nevertheless, a step forward in the outlined attempt to investigate such conditions in more detail. Two possible interpretations of the lack of latent learning demonstrated in the experimental situation of the Spence and Lippitt experiment have been ruled out and can be rejected: (1) It seems that the animals in the Spence and Lippitt experimental situation really did *not* learn the location of the food. If fear alone were the cause for their avoidance of the food-box, then the animals in our Group I, who had no cause to fear either of the two goal-boxes more than the other, should have run to the food side when made hungry. (2) Furthermore, it appears that the presence of the appropriate goal-object in the Spence and Lippitt situation was also not a sufficient factor to account for the absence of latent learning found by them. For again, if this strong positive valence were responsible for narrowing the animal's field and thus disabling it from perceiving the food, then the animals of our Group I should have learned the location of food in the oridinal training when the non-food side was empty.

It is thus impossible to account for the results obtained by Spence and Lippitt in terms of either of the two hypotheses advanced above. This

narrows down the search for the responsible factors. At least two further possibilities remain as yet untested, and might form the basis of later hypotheses concerning the conditions that inhibit or facilitate the occurrence of latent learning.

1. The degree of motivation of the animal during the original training situation may be a determining factor. In all of the situations so far yielding negative results on latent learning, the animals were under a high degree of motivation, inappropriate to the goal-object the location of which was to be learned. One might suggest that the threshold for perceiving an inappropriate goal-object rises with the increase in an unrelated drive. The results of an experiment formed also by Spence and Lippitt, but as yet published only in abstract (5), as well as those of another performed by Meehl and MacCorquodale (3), make this a reasonably plausible hypothesis. Using an essentially similar set-up as was employed in the first Spence and Lippitt experiment, latent learning was obtained when the animals were run completely satiated for both food and water. On the other hand, some preliminary, unpublished experiments performed in this laboratory give rather contradictory results, emphasizing the need for further research.

2. A second interpretation would suggest that rats are incapable of learning the location of goal-objects that are inappropriate to their then-and-there ruling drive. This would assume that rats, unless specifically trained to do so, have difficulty in discriminating their drives and the goal-objects appropriate to them. If this were true, then all they would be able to learn would be discriminanda—this side is 'black,' this is 'white,' etc.—and general 'goodness'—this side is 'good,' this other 'bad.' Perhaps they have difficulty in discriminating between the goodness of water and that of food, and do not ordinarily know that water is good when 'one' is thirsty, and food, when 'one' hasn't been fed. If such were the case, it would be expected that all further latent learning experiments, using rats which have not been trained specifically to discriminate their drives, and using an experimental situation in which what must be remembered is the location of an inappropriate goal-object, would fail in demonstrating latent learning.

Summary

1. An experiment was carried out to test two possible interpretations of the absence of latent learning in an experiment performed by Spence and Lippitt (4) which showed that rats run thirsty in a simple Y- or T-maze containing water on one side and food on the other did not shift to the food side when subsequently deprived of food:

a. The animals really knew the location of the food, but continued running to the water side because of a fear reaction induced on the food side.

b. The animals did *not* learn the location of the food. The inhibition of latent learning was caused by the presence of a strongly valenced positive goal-object—water—which narrowed the interest and the 'cognitive map' of the animal.

2. To test these hypotheses, two groups of 12 rats each were run in simple alley maze of Y-shape. The Experimental Group was run with food in one end-box, and neither food nor water in the other; the Control Group was run with food in one and water in the other. Both groups were run first under 22 hours of thirst motivation, over a period of five days, with one free run and three forced runs per day. On the seventh and eighth calendar day both groups were run under 22 hours of food deprivation for a total of six free runs per day.

If either of the two suggested hypotheses were true, the Experimental Group should demonstrate latent learning.

3. All of the animals in the Control Group continued to run to the water side on the first critical trial under hunger motivation. This again confirms the results of Spence and Lippitt (4).

4. While significantly more animals in the Experimental Group chose the food side on the first critical trial than on the free trial of the fifth day under thirst motivation, it is reasonable to suppose that this difference was due to a fluctuation on the fifth day rather than to latent learning during the training situation. This is particularly likely, inasmuch as significantly fewer animals chose that side on the fifth day than on the fourth. Furthermore, the performance on the first critical trial is not significantly different from that on the free trial of the fourth, third, second, or first day under water deprivation.

5. No significant difference was found between the performance of the Experimental Group and that of the Control Group on the 11 trials under food deprivation following the first critical trial.

6. In the light of these results, the two suggested hypotheses have to be rejected.

(Manuscript received April 25, 1949)

References

1. Kendler, H. H. An investigation of latent learning in a T maze. *J. comp. physiol. Psychol.*, 1947, 40, 265–270.
2. Kendler, H. H., & Mencher, H. C. The ability of rats to learn the location of food when motivated by thirst—an experimental reply to Leeper. *J. exp. Psychol.*, 1948, 38, 82–88.
3. Mehl, P. E., & MacCorquodale, K. A further study of latent learning in the T maze. *J. comp. physiol. Psychol.*, 1948, 41, 372–396.
4. Spence, K. W., & Lippitt, R. O. An experimental test of the sign-gestalt theory of trial and error learning. *J. exp. Psychol.*, 1946, 36, 491–502.
5. Spence, K. W., & Lippitt, R. O. 'Latent learning' of a simple maze problem with relevant needs satiated. *Psychol. Bull.*, 1940, 37, 429.
6. Tolman, E. C. *Purposive behavior in animals and men.* New York: Appleton-Century, 1932.
7. Tolman, E. C. Cognitive maps in rats and men. *Psychol. Rev.*, 1948, 55, 189–208.
8. Tolman, E. C., & Gleitman, H. *J. exp. Psychol.* (In press)

ACQUIRED DISTINCTIVENESS OF CUES: II. SELECTIVE ASSOCIATION IN A CONSTANT STIMULUS SITUATION[1]

BY DOUGLAS H. LAWRENCE

Stanford University

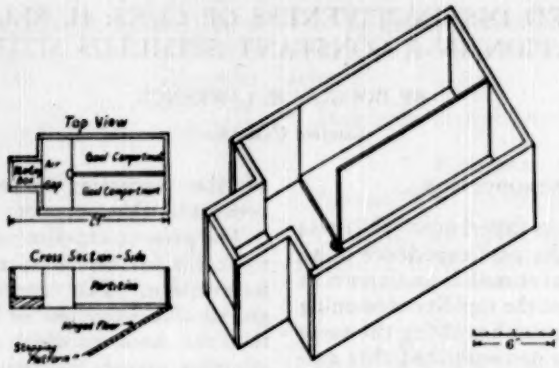
Introduction

In a previous experiment (7), it was shown that the past experience of an *S* with various stimuli is an important determinant of the rapidity of learning in a new situation involving the same cues. It was demonstrated that animals who had learned to respond to a given set of cues in a simultaneous discrimination were facilitated in learning a successive discrimination employing the same cues, whereas animals who had learned to 'ignore' these cues in the simultaneous discrimination were retarded. The design of the experiment ruled out explanations of these transfer effects in terms of the generalization of the same instrumental behavior from the first to the second task as well as the possible differential effects of acquired reward, acquired drive and like factors. The explanation offered was in terms of the 'acquired distinctiveness' of cues. It was postulated that during discrimination learning a mediating process is established that functionally alters the relative discriminability of the stimuli in the learning situation. Consequently when these same cues appear in a second learning situation, the rapidity with which they become associated with instrumental behavior is partially determined by the nature of the mediating process that is present on that occasion.

The present experiment is designed to test a further implication of this formulation. The concept of acquired distinctiveness of cues implies that the *extent* to which each of two stimulus aspects becomes associated with a response is dependent in part on the *S*'s previous experience with those stimuli. Equal physical opportunity to become associated does not ensure equal strengths of association. More concretely, this means that even though a discrimination situation contains two stimuli, both of which are relevant, they need not become equally associated with the instrumental behavior. If, through previous experience, one of the two has acquired distinctiveness whereas the other has not, the former will tend to become more strongly associated with the instrumental behavior than will the latter. Thus through proper training it should be possible to modify the extent to which any given stimulus aspect is a determinant of behavior in a new situation.

This problem is of importance in relation to the controversy between continuous and non-continuous theories of discrimination learning (2, 4, 5, 8, 9, 13, 14). If the continuity theory is interpreted literally, it would seem to imply that all aspects of the stimulus situation should become equally associated with the instrumental behavior, except for inherent differences in the relative difficulties of the various stimulus aspects. These differences are characteristic of the stimulus aspect *per se* and are

[1] The present paper is a portion of a dissertation presented to the Faculty of the Graduate School of Yale University in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The author is indebted to Dr. Neal E. Miller for suggestions and criticisms in the planning, execution and interpretation of this experiment. He also wishes to thank Dr. Clark L. Hull and Dr. Lloyd H. Beck.



STIMULTANEOUS DISCRIMINATION APPARATUS

FIG. 1. The goal compartments of this apparatus were lined with gray cardboard of different brightnesses. These linings could be interchanged between compartments. The floor was hinged so that an animal could be punished on wrong choices by being dropped onto a platform below. An air gap separated the starting box from the goal compartments, thus reducing any tendency the animal might have to rush blindly into the nearest compartment. On the far side of the gap was a small semicircular platform on which the animal stepped, thereby ensuring that each choice was made from a position directly in front of the division between the goal boxes. The floor of the starting box was raised in order to enhance the visibility of the stimuli. The rear wall of the starting box was moveable and the animal could be forced to make a choice in a specified period of time. This apparatus was used both as a successive and a simultaneous discrimination.

independent of the previous experience of the animal. On the other hand if the non-continuity theory is taken literally, it would seem to imply that one and only one of the possible stimulus aspects becomes associated with the instrumental behavior. The present formulation mediates these two positions by predicting that the degree to which the various aspects become associated with the response is partially determined by the previous experience of the *S* with those aspects, but that this is a relative rather than an all-or-none affair.

The plan of this experiment is to train one group of animals to respond on the basis of a given cue in a successive discrimination and to ignore a second cue that is present during learning; the other group of animals learns the converse discrimination. The two groups are then trained in a simultaneous discrimination where both of these cues are relevant, i.e., the animals can respond correctly on the basis of either one or both.[3] At the completion of learning, the animals are then tested to determine the extent to which the dependency of the instrumental response on each of the cues has been modified by the previous training with those cues.

[3] As emphasized in a previous experiment (7), training on a successive discrimination and testing on a simultaneous discrimination rules out any explanation of the resultant transfer effects in terms of the generalization of overt, instrumental responses from the training to the testing situation. In a successive discrimination, the animal is always rewarded for making the same response (such as turning left) in the presence of a given cue; in the simultaneous discrimination he must half the time turn left and half the time turn right in the presence of this same cue. Consequently any transfer effects must be explained by some other mechanism than the generalization of the habit strength resulting from a specific association between the cue and the overt, instrumental behavior.

Subjects and Apparatus

The *S*s in this experiment were 45 male albino rats approximately 60 days old at the beginning of training. Five of these were discarded during adaptation training either because of unusual timidity or because of excessive weight loss due to the restricted diet.

The discrimination apparatus used in this experiment is shown in Fig. 1. It has been more fully described in a previous paper (7). It consisted of a starting box that faced directly into two rectangular goal compartments that were side by side. Separating the starting box from the goal compartments was a 4½ in. air gap that the animal had to step across. In this way the rat's tendency to rush into the nearest compartment was reduced. A semicircular landing platform, directly in front of the dividing partition between the two goal compartments, ensured that the animal would always be in a position such that both compartments were visible to him while making a choice. The floor of the apparatus was hinged so that the experimenter could release it whenever the animal made an error. When this occurred the animal fell onto a platform, 2½ ft. below. The rear wall of the starting box was movable so that the animal could be forced to make a choice if it hesitated more than 30 sec. after being placed in the apparatus. An animal was never forced, however, once it had placed its front feet on the landing platform.

Two sets of cues were employed in this experiment, black-white and the presence or absence of chain curtains. The black-white cues consisted of painted cardboard strips that were used to line the walls and floor of the goal compartment. These could be readily interchanged between compartments. The chain curtains were made by extending small brass rods across the top of a compartment, and suspending from these at inch intervals a series of silver colored, lamp-socket pull chains. The bottom ends of the chains rested on the floor so as to reduce any tendency to swing back and forth. Two chain curtains were present in a given compartment at any one time; one at the very entrance to the compartment and one 4 in. in front of the food cup. In this way the chains were not only a visual cue for the animal, but he was tactually stimulated by them at every point in the goal compartments.

This apparatus was used both as a successive discrimination and as a simultaneous discrimination. Consequently, when an animal was shifted from one type of training to the other, there was very little opportunity for external inhibition to disrupt performance. The only change made in the stimulus situation was the change in the configuration of the stimuli.

The entire experiment was run inside a white, cardboard dome that has previously been described (7). The lighting was entirely indirect. In this way extraneous cues, such as shadows and the like, were held at a minimum.

Adaptation Training

Prior to the experiment all animals were tamed by being made thirsty and then being allowed to drink while held by the experimenter. Following this they were put on a 24-hr. feeding schedule. For two hours a day they had all the dry Purina Laboratory Chow they would eat without water. The food was then taken away and the water bottles returned to the cages. This feeding schedule was continued throughout the experiment. The animals were given 20 forced trials on the apparatus in order to train them to step across the air gap. The width of the gap was gradually increased during these trials until by the 11th it was the full 4½ in. During these trials the four possible stimulus combinations of black-white and chains and no-chains were equally often present. Half the trials were to the right compartment and half to the left. On each trial the animal was allowed to eat a small pellet of moist food in the food cup.

The following day the animals were given a 'dry run' of 10 trials with the various combinations of cues again present. This dry run consisted of 10 free choices on which the experimenter arbitrarily decided that either the right or left compartment would be 'correct' on that trial. If an animal made an incorrect choice, he was dropped onto the platform, immediately picked up, returned to the starting box, and the trial continued until he chose correctly. If he refused to leave the starting box for longer than 30 sec. on any choice, he was forced by advancing the rear wall of the starting box. By means of this dry run most animals tended to get over their initial startle reactions to being dropped or pushed; it also tended to break down position habits.

Experimental Design

A. Preliminary Training on the Successive Discrimination

At the completion of adaptation training, the 40 animals were divided randomly into two groups of 20. Both of these then were trained on the *successive* discrimination in order to set up differential habits of response to the two stimulus aspects. One group (the black-white group) learned to respond to the black-white cue and to

ignore the presence or absence of chains. The other group (the chain group) learned to respond to chains and to ignore the presence of black or white. This was accomplished in the following way.

On half the trials both goal compartments of the apparatus were black; on the other half of the trials they were white. On half the black trials the chain curtains were present in both alleys and on the other half the chain curtains were absent. The same was true on the white trials. The black-white group had to learn to choose the goal compartment on one side of the apparatus (the right or the left) when black was present and to choose the other goal compartment (the left or the right) when white was present, irrespective of the presence or absence of the chains. They were responding to black-white and ignoring the chains, no-chains. Half of these animals learned to turn right to black and left to white, and the other half learned the converse pattern of responses.

The chain group, on the other hand, learned to choose one goal compartment when the chain curtains were present and the other goal compartment when there were no chains, irrespective of whether or not the compartments were black or white. Thus, they were responding to the presence or absence of chains and ignoring black-white. The cue that each group learned to respond to will hereafter be called the 'preferred' cue of that group, and the cue they learned to ignore will be called the 'non-preferred' cue.

The training procedure was the same as that employed during the dry run in adaptation training, i.e., a trial consisted of one correct choice. Both groups were given 80 trials, 10 trials a day, on this successive discrimination by the end of which time they both were responding better than ninety percent correctly.

B. Training on the Simultaneous Discrimination

After having been trained to respond to one set of cues and to ignore the other in the successive discrimination, both groups were then trained on exactly the same *simultaneous* discrimination. In this way the previous training of responding differentially to the two sets of cues had an opportunity to determine the extent to which the various stimulus aspects of the new situation were associated with the new instrumental responses. This simultaneous discrimination consisted of having one goal compartment of the apparatus white and the other black on each trial. Chain curtains were always present in the black compartment and always absent in the white compartment for half the animals. The relation between black-white and chains and no-chains was reversed for the other half of the animals. These compartments were interchanged in an irregular order from side to side on succeeding trials.

The design was so arranged that a fourth of the animals in each group was trained to each of the four possible stimulus combinations as the positive cue. The previous experience of the animals, such as whether or not they had turned right or left to black on the successive discrimination, was balanced. During the first 10 trials of this simultaneous training, the animals were forced out of the starting box if they had not made a choice by the end of 30 sec. From the 11th trial on, however, they were never forced but were permitted all the time necessary to make a choice. Training throughout was by the correction method. All animals were given 30 trials, 10 trials a day. During the last ten of these they were responding almost perfectly.

C. Test Number 1: Opposition of Cues

As a test of the dependence of the instrumental response on each of the two cues, two test trials were inserted in the last ten simultaneous training trials. On these test trials the two cues were opposed, i.e., if the animal had been responding to black-chains as the positive configuration of cues and white-no chains as the negative configuration, he was presented on the test trials with the two new configurations of black-no chains and white-chains. Thus, the positive aspects of the two cues were opposed; black was in one compartment and chains in the other. If the animal chose the black it would indicate that he was responding primarily in terms of black-white; if he chose the chains, he was responding primarily to chains and no-chains. The test trials were the 24th and the 29th of the training series. On the first test, the preferred cue remained on the same side of the apparatus as it had been on the preceding trial for half the animals, and was on the opposite side for the other half. The converse was true on the second test trial. The sequence of responses preceding the first test was right, right, left; and the sequence preceding the second was left, left, right. The animals were rewarded for whichever choice they made on these test trials.

D. Test Number 2: Relearning on a Single Cue

As a second test of the extent to which each of the two cues was a determinant of the instrumental response, half the animals (10) of the black-white group and half the animals (10) of the chain group were retrained on the simultaneous discrimination with only one of the two cues

present. Half of these black-white group animals were retrained on their preferred cue first and half on their non-preferred cue. The same was true of the chain group animals. At the completion of this training, all animals were retrained on the opposite cue.

When chains and no-chains was the cue, both compartments were a neutral gray; thus, an 'unfamiliar' aspect was present in the stimulus situation. When black-white was the cue, the chain curtain was removed leaving both compartments in a 'no-chains' condition. This was a 'familiar' situation for the animals. Consequently, there was a differential effect introduced into this test. The influence of this differential effect on performance, however, should show up as a difference in the difficulty of the black-white and chain and no-chain cues. It would be balanced out as far as the performance on the preferred and non-preferred cues was concerned. By means of this test it was possible to determine if the animals relearned the discrimination more rapidly on the preferred cue than they did on the non-preferred cue. Training was given by the usual correction method. Thirty trials, 10 trials a day, were given on each cue.

E. Test Number 3: Reversal of Learning on a Single Cue

The other half of the black-white group and the other half of the chain group also were retrained on the simultaneous discrimination with only a single cue present. In this case, however, the positive and negative aspects of the cues were reversed. If an animal had previously been trained to black-chains as positive and white-no chains as negative, he was retrained on black-white alone but with white positive and black negative. In all other respects the design of this test was the same as the previous one. Consequently, a difference in the rates with which the animals reversed the discrimination on the preferred and non-preferred cues would indicate the relative extent to which these cues were determining the animal's behavior. All animals were given 40 trials, 10 trials a day, on these reversed discriminations.

Results

A and B: Training on the Successive and Simultaneous Discriminations

The learning curves for the black-white group and the chain group on the successive discrimination and the simultaneous discrimination are shown in Fig. 2. It is apparent that

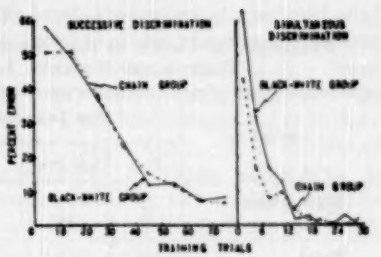


Fig. 2. Each point on these curves represents 20 animals. The curves on the left represent the learning of these animals on a successive discrimination in which there were two cues, a relevant one and an irrelevant one. The relevant one was the cue that the animals had to respond to in order to learn the discrimination. The irrelevant one bore no relationship to the rewards or punishments in the situation and had to be 'ignored.' The Chain group had the chain curtains as a relevant cue, the Black-white group had brightness as the relevant cue. After this training, all animals were trained on a simultaneous discrimination in which both of the former cues were now relevant. The learning on that discrimination is represented by the curves on the right. It is apparent that there was no difference in the rate at which the two groups learned the two discriminations.

the two *successive* discriminations were unusually well matched as regards difficulty of learning. The animals learned to respond to black-white and to ignore chains and no-chains as readily as they learned the opposite behavior. The learning curves for the *simultaneous* discrimination indicate that this difference in the cue responded to by the two groups had little if any influence on the rate with which they learned the simultaneous discrimination problem where both sets of cues were relevant. There is some indication that the chain group began to learn at a slightly faster rate than did the black-white group, but this difference is slight. The other important feature of these curves is that after the 15th trial on the simultaneous discrimination both groups were responding with better than 95 per-

TABLE I

FREQUENCY OF CHOICE OF THE PREFERRED AND NON-PREFERRED CUES WHEN THE POSITIVE AND NEGATIVE ASPECTS OF THE TWO ARE OPPOSED

Group	First Test (24th Trial)				Second Test (29th Trial)			
	Pref.	Non-Pref.	*df*	Chi-sq.	Pref.	Non-Pref.	*df*	Chi-sq
Black-white	17	3	1	9.8	18	2	1	12.8
Chain	10	10	1	0.0	10	10	1	0.0
Total	27	13	2	9.8	28	12	2	12.8
Pooled chi-sq.			1	4.9			1	6.4
Interaction			1	4.9			1	6.4

(Chi-square with 1 *df:* 3.84 at 5%; 5.41 at 2%; and 6.64 at 1%)

cent accuracy. Thus the opposition test was introduced after the discrimination had been thoroughly mastered.

C. Test Number 1: Opposition of Cues

The first test of the relative dependence of the instrumental responses on the preferred and the non-preferred cues was obtained by opposing these two cues on the 24th and the 29th trials of the simultaneous discrimination. If the animal had been responding to black-chains as positive and white-no chains as negative, he was now given a choice of black-no chains and white-chains; the positive aspect of the preferred cue was combined with the negative aspect of the non-preferred cue and conversely. The results of these two test trials are shown in Table I.

Keeping in mind that the black-white cue is the preferred cue for the black-white group but the non-preferred cue for the chain group, it is apparent that the choices in Table I were determined primarily by two factors: a tendency to choose on the basis of the preferred rather than the non-preferred cue, and a tendency to choose on the basis of the black-white rather than the chain cue. Consequently, the black-white animals all tended to choose the preferred cue because the two factors were working in the same direction, whereas the chain animals split their choices because the factors were working in opposite directions. This is shown by the pooled chi-square and the interaction chi-square.

The pooled chi-square is based on the distribution of all 40 choices between the preferred and non-preferred cues. On the 24th trial this chi-square value is significant between the two and five percent levels of confidence, and on the 29th trial, between the one and two percent levels. Thus, while these two values are not independent in that they are based on the same animals, the two taken together give ample support to the hypothesis that the preferred cue tends to be a more important determiner of the instrumental response than does the non-preferred cue. Similarly, the two interaction chi-squares, which are significant between the two and five and between the one and two percent levels of confidence on the 24th and 29th trials, respectively, show that the black-white cue tends to be chosen over the chains and no-chains cue.

None of these chi-square values has been corrected for continuity because

the addition theorem was used. If the pooled chi-squares are corrected for continuity, the values for the first and second tests remain within the same levels of confidence as before. It can be concluded, therefore, that the original training in responding to different aspects of the stimulus configuration influenced the extent to which these aspects determined the instrumental response. The other determining factor was the 'natural dominance' of black-white over chains and no-chains in this situation.

D. Test Number 2: Relearning on Single Cues

The second test of this hypothesis consisted of retraining half the animals on the simultaneous discrimination when only one of the two sets of cues was present. When this retraining was completed, the animals were then retrained on the other cue. The positive and negative aspects of the cues remained the same as in the original training. The relearning curves on the preferred and non-preferred cues are shown in the left-hand panel of Fig. 3.

The curves indicate that these animals show more retention and possibly faster relearning on the preferred cue than they do on the non-preferred one. All other factors are balanced so this differential effect must be attributed to the fact that the animals have been taught to respond to the preferred cue and to ignore the non-preferred one.

The data on which these relearning curves are based are shown in Table II. This table indicates that the same two factors determine the relearning score that determined the choice responses in the previous opposition test. As can be seen from the overall averages, the animals tend to make fewer average errors on the discrimination employing the preferred cue (2.7) than on the one employing the non-preferred cue (4.5). Similarly, they tend to make fewer average errors on the black-white cue (1.6)than on the chains (5.6). The statistical significance of these differences is demonstrated by the analysis of variance of the transformed scores in the left half of Table III.

The design of this relearning test involves three different variables, each in two degrees: preferred versus non-preferred cues, first versus second relearning session, and black-white versus chains, no-chains. If a given group of animals first relearned the

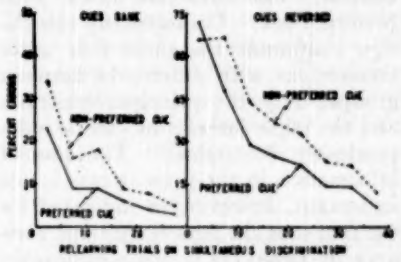


FIG. 3. The curves on the left are for the same 20 animals under two test conditions. The curves on the right are for a different group of 20 animals under a different set of test conditions. All animals had been trained to respond to one cue and to ignore a second in a successive discrimination in which these two cues were present. They were then trained on a simultaneous discrimination in which these two cues were both relevant, i.e., the discrimination could be learned in terms of either one or both. As a test of which of the two had become more strongly associated with the instrumental behavior, one cue was removed and training continued. The solid line in the left panel shows the performance of these animals when the cue present in the apparatus is the one they have been taught to respond to; the dotted line, the cue they have been taught to ignore. Exactly the same is true of the animals on the right except that they had to reverse their habit on the single cues. The fact that the preferred cue gives the faster learning in each instance indicates that training to respond to a cue in the successive discrimination influences the rate of relearning and reversing a habit in the simultaneous discrimination.

TABLE II

MEAN NUMBER OF INITIAL ERRORS FOR 30 TRIALS DURING THE RELEARNING OF THE SIMULTANEOUS DISCRIMINATION WITH SINGLE CUES
(Positive and Negative Aspects of Cues Unchanged)

Group	N	Preferred Cue	First Test		Second Test	
			Test Cue	Mean	Test Cue	Mean
1A	5	Black-white	Black-white	1.2	Chains	7.6
2A	5	Black-white	Chains	5.8	Black-white	0.6
3A	5	Chains	Black-white	2.2	Chains	5.8
4A	5	Chains	Chains	3.2	Black-white	2.4

Overall Averages: (Preferred 2.7, Non-preferred 4.5); (Black-White 1.6, Chains 5.6)

black-white discrimination as the preferred one, then of necessity on the second relearning they had the chains, no-chains discrimination as the non-preferred one. Consequently this design confounds the three first order interactions with differences between groups; only the principal contrasts and the triple interaction can be independently determined. The loss of information in the present case is not important, however, as indicated by the fact that the between groups variance in Table III is not significantly different from the error term.

The difference between the preferred and non-preferred cues is significant between the one and five percent levels of confidence. The difference between the black-white and chains is significant at well beyond the one percent level of confidence. Thus the present relearning test substantiates the findings of the opposition test.

E. Test Number 3: Reversal of Learning to Single Cues

In the first test it was shown that an animal tends to choose the preferred over the non-preferred cue; in the second test it was shown that relearning on the preferred cue is faster than on the non-preferred. A third

TABLE III

ANALYSIS OF VARIANCE OF THE INITIAL ERROR SCORES ON THE RELEARNING AND REVERSAL TESTS WITH SINGLE CUES. ERROR SCORES TRANSFORMED TO CORRECT FOR CORRELATION BETWEEN MEANS AND VARIANCES (11). EXPERIMENTAL DESIGN RESULTED IN THE CONFOUNDING OF FIRST-ORDER INTERACTIONS WITH BETWEEN-GROUPS VARIANCE

Item	Comparison	df	Relearning Test		Reversal Test	
			Variance	F	Variance	F
1	Between-groups and confounded interactions	3	56.13	—	54.81	1.14
2	First vs. second test	1	21.46	—	21.62	—
3	Black-white vs. chains	1	1875.53	29.10**	582.93	12.10**
4	Preferred vs. non-preferred	1	421.85	6.55*	371.49	7.71*
5	Triple interaction	1	32.94	—	21.62	—
6	Between individuals treated alike	16	74.59	1.16	53.43	1.11
7	Error	16	64.44		48.17	

* Significant between the one and five percent levels of confidence.
** Significant beyond the one percent level of confidence.

TABLE IV

MEAN NUMBER OF INITIAL ERRORS FOR 40 TRIALS DURING THE REVERSAL OF THE SIMULTANEOUS DISCRIMINATION WITH SINGLE CUES
(Positive and Negative Aspects of Cues Reversed)

Group	N	Preferred Cue	First Test		Second Test	
			Test Cue	Mean	Test Cue	Mean
1B	5	Black-white	Black-white	11.4	Chains	19.2
2B	5	Black-white	Chains	16.4	Black-white	7.2
3B	5	Chains	Black-white	12.6	Chains	13.6
4B	5	Chains	Chains	14.0	Black-white	12.6

Overall Averages: (Preferred 11.6, Non-preferred 15.2); (Black-white 11.0, Chains 15.8)

test of the relative dominance of the preferred and non-preferred cues was made by having the animals relearn the discrimination on each cue separately but with the positive and negative aspects of these two cues reversed. Any difference in the rate with which these reversals are established must be attributed to the influence of previous training in learning to respond to the preferred cue. This test was made with the 20 animals that had *not* been used in the previous test.

The relearning curves of these animals on the preferred and non-preferred cues during the reversal of the discrimination are shown in the right-hand panel of Fig. 3. It is apparent that reversal is faster on the preferred cue than it is on the non-preferred. The data on which these curves are based are summarized in Table IV. Again it appears that the same two factors are determining the reversal rate in this test that determined the differential effects in the other two tests.

The overall averages indicate that the reversal of the preferred cue results in fewer average errors (11.6) than does the reversal of the non-preferred cue (15.2), and that the reversal of the black-white cue results in fewer average errors (11.0) than does the reversal of the chain cue (15.8). The significance of these differences is shown in the analysis of variance of the transformed scores given in Table III. The difference between the preferred and non-preferred cues is significant between the one and five per cent levels of confidence. The difference between the black-white and chain cues is significant beyond the one percent level of confidence.

DISCUSSION

In a previous experiment (7) it was shown that if an animal is first trained to respond to a given cue, that cue is more readily associated with a new instrumental response than it is without such training. The present experiment has increased the generality of the previous findings in several ways. The introduction of a new stimulus dimension, the presence or absence of chain curtains, indicates that the results are characteristic of all cues of the type employed in these two experiments, i.e., black-white, wide-narrow compartments, rough-smooth floors, and chain curtains. This is especially important in that these stimuli were chosen because they ensure that the animal will be physically stimulated by them on every choice; they are what Spence (13) calls perceptually simple cues.

Consequently the role of 'acquired distinctiveness of cues' in learning such discriminations should be less apparent than in more complex stimulus situations. The fact that it is demonstrated consistently on this type of cue underlines the importance of this phenomenon.[3] Similarly, the employment in this experiment of three new techniques of measurement (the opposition test, the relearning test and the reversal test) increases the confidence with which conclusions can be drawn.

More importantly, the present experiment has shown that acquired distinctiveness of cues has a more general influence on performance than could be inferred from the previous experiment. In the previous experiment, the test of this factor was in terms of the rate with which new instrumental responses were learned; the emphasis was on the intial stages of the learning. In the present experiment, all the tests were introduced after the animals were performing perfectly on the simultaneous discrimination. Nevertheless, the opposition test indicates that the animals tended to choose the preferred cue, i.e., the one they originally had been taught to respond to, rather than the non-preferred cue. Similarly, the relearning test indicates that the animals show more retention on the preferred cue than on the non-preferred. It is apparent, therefore, that this phenomenon is not limited to the initial stages of learning, but that it continues to be an important determinant of performance even when learning is far advanced.

Another aspect emphasized by the present experiment is the selective function of previous familiarity with a cue on the learning that takes place in a new situation. In the previous experiment familiarity was shown to function in this way when there was a *relevant* and an *irrelevant* cue present. By a relevant cue is meant one that is correlated with the rewards and punishments in the situation and to which the animal must respond in order to learn the discrimination. By an irrelevant cue is meant one that is present on each trial but that has no correlation with the rewards and punishments. In the present experiment both the preferred and non-preferred cues were relevant during the training on the simultaneous discrimination. The discrimination could be learned on the basis of either one or both; there was nothing in the physical situation or the training procedure that favored one cue over the other. Despite this equality of opportunity to become associated, the instrumental responses proved to be more dependent on the cue that the animal previously had learned to respond to than they were on the non-preferred cue. In this sense the acquired distinctiveness of cues has a selective function that determines in part the extent to which any given aspect of the stimulus situation will become as-

[3] An explanation of these results in terms of orientation behavior may be possible. Such behavior can never be ruled out of a learning situation; it may be present though unobserved. The likelihood of this type of explanation is small, however, because the apparatus and stimulus dimensions were chosen to minimize the need for such behavior. The apparatus was an enclosed box so as to reduce extraneous cues, the two relevant stimuli were side by side without spatial separation, and the two compartments were visible simultaneously to the animal both from the starting box and the point of choice. The cues, black-white and chains, no-chains, were of such a nature that the animal was certain to be physically stimulated by them on each choice. If the explanation of these results is in terms of orientation behavior, this behavior must be so prevalent and so fundamental in discrimination learning that it would only serve to emphasize the importance of the present results.

sociated with the instrumental behavior.

An acceptable explanation of these phenomena must not only account for the findings mentioned so far but also for the apparently anomalous results on the reversal test. In that test it was found that animals were able to reverse the discrimination on the preferred cue alone significantly faster than they were on the non-preferred cue. But the other two tests, the opposition test and the relearning test, indicate that the preferred cue was a stronger determinant of the instrumental behavior than was the non-preferred cue. It would appear, therefore, that the instrumental responses were more strongly associated with the former than they were with the latter cue. On this assumption, the results indicate that it is easier to reverse a strong habit than it is to reverse a weak one. Such a conclusion would be in contradiction to previous findings in habit reversal experiments (9, 14).

This assumption of a stronger association between the instrumental responses and the preferred cue is justified only if the relationship between the stimulus and the instrumental response is conceived of as a unitary one that depends entirely on the strength of association that has been established between them. In such a case performance is a direct measure of the strength of association. If on the other hand this relationship is conceived of as essentially a two stage process, the more rapid reversal on the preferred cue is not necessarily contradictory to a stimulus-response interpretation of learning.

A behavioral conception of discrimination learning that takes these two stages into account was outlined in the previous experiment in terms of the acquired distinctiveness of cues. The first stage of such learning was postulated to be the establishment of a mediating process that functionally altered the discriminability of the relevant cues thus making them more distinctive. The second stage consisted of the establishment of an association between this more distinctive stimulus pattern and the instrumental responses. The mediating process was assumed to be set up as the result of previous training on these cues and to transfer to similar situations. As a consequence the learning of new instrumental responses would be facilitated.

This formulation is able to account for the results on the reversal test as well as the other findings. What this formulation implies is that through previous experience two sets of cues that initially are equally hard to discriminate can be modified to such an extent that one set becomes relatively easier than the other set. Subsequently these two sets of cues would have the same influence on behavior as would two sets of cues that were of different degrees of discriminability from the very beginning of learning. An example of this latter type is contained in the present experiment. In the opposition and relearning tests, the black-white was a more important determinant of performance than was the chains, no-chains. It would appear, therefore, that the instrumental behavior was more strongly associated with the black-white than with the chains. Nevertheless the rate of reversal was faster on black-white than it was on the chains. This parallels the findings on the preferred and non-preferred cues. In the case of the black-white, however, these two results can be accounted for by stimulus-response principles. Two factors are involved: (a) in that black-white is more discriminable than chains, no-chains, i.e., there is less generalization of habit strength between the positive and negative aspects of the former than between the latter, each reward and punishment will result in a greater differential habit strength on black-white than on chains, no-chains, and (b) at the completion of training the absolute differential on black-white will be greater

than it is on the chains. This latter factor accounts for the finding that black-white tends to dominate over chains in the opposition and relearning tests, and the other factor accounts for the faster reversal on the black-white discrimination. If the same logic is applied to the preferred and non-preferred cues, it is apparent that the findings on the reversal test are in no sense anomolous or unexpected, but follow directly from the hypothesis of acquired distinctiveness of cues.

This formulation of discrimination learning in terms of the acquired distinctiveness of cues is in many ways analogous to Lashley's theory (5, 6). Both formulations conceive of the relationship between the stimulus situation and the instrumental response as dependent on at least two variables, a mediating process and an association process. (According to Lashley, the former is a matter of attention.) Both formulations are capable of accounting for the results in this experiment and the previous one. The Lashley type or non-continuity theory, however, is usually contrasted with the stimulus-response type or continuity theory. But the fact that the present formulation is an extension of concepts to be found in stimulus-response theories of learning (3, 10) indicates that the phenomena of attention are not a crucial distinction between the continuity and non-continuity theories. On the other hand, the acceptance of the present formulation does require a change in emphasis of strict stimulus-response interpretations of discrimination learning (1, 3, 12).

The hypothesis of the acquired distinctiveness of cues and the non-continuity theories of learning differ as to the characteristics they ascribe to the mediating process. The latter tend to emphasize the all-or-none characteristic of shifts in attention during the presolution period of learning. The present formulation assumes that the changes in the mediating process are of a gradual and continuous nature. Both agree, however, that the end result of such modifications may be a qualitative change in the 'perceptual' characteristics of the situation. Because of this difference in emphasis, the two formulations tend to disagree also as to the extent that attention limits the association between various stimulus aspects and the instrumental behavior. The non-continuity theories tend to emphasize that only the aspect attended to becomes associated with the stimulus. The present formulation assumes that the relevant aspect of the stimulus becomes relatively more distinctive than the other aspects; there is a reordering of the 'association values' of the various aspects but none is necessarily excluded from the association.

The present experiment contains some evidence on this latter point. In both the relearning and the reversal tests the animals showed some dependence on the non-preferred cue. In neither of them did learning start at the 50 percent level of accuracy as would be expected if the animals had made no associations with the non-preferred cue. Thus, in the present situation, at least, both cues were associated with the instrumental responses even though the animals had been trained to respond to one cue and to ignore the other.[4] The implication of this result is that in this type of experimental situation the mediating process modifies the readiness with which any given aspect of the stimulus situation becomes associated with the instrumental behavior rather than excluding all aspects but one. This can be accepted readily by non-continuity theories. Their reliance on the exclusion principle is primarily one of emphasis rather than of theoretical necessity.

Summary

Twenty animals were trained on a successive discrimination to respond to

[4] This evidence is by no means crucial for the non-continuity theories. For them a stimulus dimension is a configural resultant; black-white and chains may not qualify as such in the present experiment. But in view of the type of training given on the successive discrimination, where these two cues were varied independently from trial to trial, it is difficult to see why they would not tend to become isolated and unitary dimensions.

a black-white stimulus and to ignore the presence or absence of chain curtains in the goal compartments. Another 20 animals were trained on the converse habit. All animals were then trained on a simultaneous discrimination in which both cues were relevant, i.e., the discrimination could be learned on the basis of either one or both. At the completion of learning, all animals were tested with the two cues opposed. The direction of choice gave an indication as to which of the two cues had the greater influence on performance. Than half the animals in each group relearned the simultaneous discrimination with only one of the two cues present. The other half of the animals relearned the simultaneous discrimination with only one cue present, but the positive and negative aspects of this cue were reversed.

1. The oppotision test indicated that the animals tended to choose on the basis of the cue they had been taught to respond to, the 'preferred' cue, rather than on the basis of the one they had been taught to ignore, the 'non-preferred' cue.

2. The relearning test indicated that relearning was faster on the preferred cue than it was on the non-preferred cue.

3. The reversal test indicated that the animals reversed the discrimination on the preferred cue more rapidly than they did on the non-preferred cue.

4. These results, taken in conjunction with the results of a previous experiment (7), indicate that the learning of new instrumental responses is faster on a familiar cue than it is on a non-familiar one; that the influence of this familiarity is not restricted to the initial stages of learning but is an important determinant of behavior even when learning is far advanced; that it has a selective influence on which of the aspects of the relevant stimulus will become associated with the instrumental behavior; and that these results are obtainable on a variety of perceptually simple stimulus dimensions.

5. These findings are interpreted in terms of the concept of the acquired distinctiveness of cues which postulates that discrimination learning is essentially a two stage process, the first stage of which is a change in the 'perceptual' characteristics of the stimulus. The similarity between this formulation in terms of stimulus-response principles and that of non-continuity theories of learning is discussed. Some evidence is presented that there is not an all-or-none character to the stimulus aspects that become associated with the instrumental responses in the present experiment.

(Manuscript received April 26, 1949)

References

1. Gulliksen, H., & Wolfle, D. L. Theory of learning and transfer. *Psychometrika*, 1938, 3, 128–149.
2. Haire, M. Some experimental data relevant to field and associative theories of discrimination learning. *J. gen. Psychol.*, 1942, 26, 267–288.
3. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
4. Krechevsky, I. A study of the continuity of the problem-solving process. *Psychol. Rev.*, 1938, 45, 107–133.
5. Lashley, K. S. An examination of the 'continuity theory' as applied to discrimination learning. *J. gen. Psychol.*, 1942, 26, 241–265.
6. Lashley, K. S., & Wade, M. The Pavlovian theory of generalization. *Psychol. Rev.*, 1946, 53, 72–87.
7. Lawrence, D. H. Acquired distinctiveness of cues: I. Transfer between discriminations on the basis of familiarity with the stimulus. *J. exp. Psychol.*, 1949, 39, 770–784.
8. McCulloch, T. L. Comment on the formation of discrimination habits. *Psychol. Rev.*, 1939, 46, 75–87.

9. McCulloch, T. L., & Pratt, J. G. A study of the pre-solution period in weight discrimination by white rats. *J. comp. Psychol.*, 1934, 18, 271–290.
10. Miller, N. E., & Dollard, J. *Social learning and imitation.* New Haven: Yale Univ. Press, 1941.
11. Snedecor, G. W. *Statistical methods.* Ames, Iowa: The Iowa State College Press, 1946.
12. Spence, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.*, 1937, 44, 430–444.
13. Spence, K. W. Continuous versus non-continuous interpretations of discrimination learning. *Psychol. Rev.*, 1940, 47, 271–288.
14. Spence, K. W. An experimental test of the continuity and non-continuity theories of discrimination learning. *J. exp. Psychol.*, 1945, 35, 253–266.

GENERALIZATION OF SECONDARY REINFORCEMENT

BY IRVING J. SALTZMAN

Indiana University[1]

Introduction

It has been shown by Denny (1) that fewer trials are required by rats to learn a simple T-maze when the goal boxes are markedly different—one white and one black—than when the goal boxes are similar—both grey. The explanation offered to account for this difference in learning rate involves the concept of secondary reinforcement, specifically, the generalization of secondary reinforcement. Denny suggested that when two grey goal boxes are used, the secondary reinforcing properties which the positive grey goal box acquires are shared, through generalization, by the similar, negative grey goal box. The result of this generalization of reinforcing value from the positive grey box to the negative grey box is that incorrect responses, as well as correct responses, are reinforced during the learning situation. As a consequence, the learning that occurs under these conditions is slower than the learning that occurs when dissimilar goal boxes are used. When markedly different goal boxes are employed, the generalization of reinforcing value that may occur is at a minimum. Consequently, only the correct responses are reinforced and learning proceeds at a significantly faster rate.

In Denny's experiment all the rats were given preliminary training on a straight runway. As a result of this runway experience, in the writer's opinion, the goal boxes that were to be the positive goal boxes on the maze, prior to the maze trials, acquired secondary reinforcing value to a substantial degree (2). And, if the generalization hypothesis is correct, the *negative, grey* goal box also acquired reinforcing value before the beginning of the maze trials. Since the secondary reinforcing value that was generalized from the positive grey box to the negative grey box was established primarily during the preliminary training on the runway, and not during the maze experience, the question arises: If the negative goal box does not already possess reinforcing value before the maze learning begins, will the learning be slower, nevertheless, for the similar goal box group than for the dissimilar goal box group? In other words, will the generalization of secondary reinforcement that is *established during the maze learning situation itself* be sufficient to produce a difference in the learning rate when similar and dissimilar goal boxes are used? The present study was designed to answer this question by eliminating the possible influence in the learning situation of previously acquired secondary reinforcement.

Experimental Procedure

A. Subjects.—Thirty-six naive, hooded, female rats, from the Johns Hopkins Psychology Department colony, were divided into three groups of 12. The rats ranged in age from 90 to 120 days at the beginning of the experiment.

B. Apparatus.—The apparatus (Figs. 1 and 2) consisted of a closed runway, a single choice, closed U-maze, four removable goal boxes and a removable starting box. All the apparatus was made of wood and painted grey, except for two of the goal boxes, one of which was painted black, and the other, white. All the apparatus, except the white goal box, was raised on small legs. Wire mesh served as the top of each piece of apparatus. The goal boxes and the starting box

[1] This study was carried out at the Psychological Laboratory, the Johns Hopkins University.

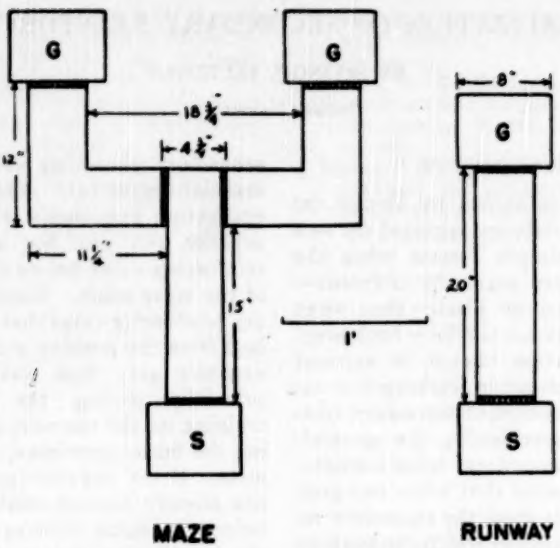


FIG. 1. Floor plan of maze and runway. S = Starting Box; G = Goal Box

were used interchangeably on both the runway and the maze.

Vertically sliding doors were located as indicated in Figs. 1 and 2 by the crosshatching. The black goal box had a small hurdle on its open side which had to be climbed by the rats in entering this box from either the maze or the runway. To enter the white goal box, which was not raised on legs, the rats had to jump down from either the maze or the runway. And to enter the grey goal box, the rats had simply to walk in. The goal boxes were not visible from the choice point in the maze.

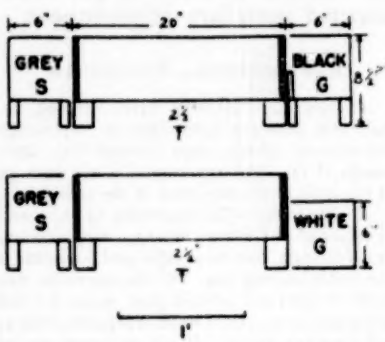


FIG. 2. Side view of runway showing position of black and white goal boxes. S = Starting Box; G = Goal Box.

C. Preliminary training.—A week before the experiment proper was begun, the rats were placed in individual cages and were put on a 23½-hour feeding schedule. The food was moist Purina (Growena) chicken mash, which was presented in small metal cups. Water was available at all times. The feeding was omitted on the seventh day and on the eighth day the runway trials were begun.

D. Runway trials.—All three groups of rats received the same treatment except that the goal box was grey for the rats in Groups I and II, and black for half the animals in Group III and white for the other half. A rat was placed in the starting box, and after a delay of 10 sec. was allowed to proceed to the goal box. After the rat entered the goal box, it was given a small cup of food from which it was allowed to eat for two min. The rat was then removed to its home cage. Five such trials were given each rat. Following the fifth runway trial, each of the rats was allowed to eat in the goal box for 20 min. In this manner the rats received their daily rations of food in their goal boxes.

This procedure was repeated for five consecutive days, at the end of which time each rat had received 25 reinforced runs to its particular goal box—grey for the rats in the first two groups, and either black or white for the rats in the third group. On the last day of runway experience,

between the second and third trials for the day, the position habits of the rats were tested on the maze. A choice of one side two out of three times was the measure employed.

E. Maze learning trials.—The learning trials on the maze were conducted on the day following the last day of runway trials. All three groups of rats received the same treatment, but the goal boxes were different for the different groups in order to control for the influence on the maze learning of already established secondary reinforcement. The rats in Group I, which had run to the grey goal box on the runway, ran to dissimilar goal boxes on the maze, one white and the other black. For half of these animals, the black goal box was positive and the white box, negative, and for the other half, the reverse was true. The positive goal box was always located on the non-preferred side of the maze. For this group, the influence on the maze learning of previously established reinforcing value was presumably eliminated. The rats in Group II, which also ran to a grey goal box on the runway, ran to two similar grey goal boxes on the maze. The positive goal box was located on the non-preferred side of the maze. For this group, the influence of previously established reinforcing value was not eliminated since the goal boxes on the maze were similar to the goal box on the runway. And for the rats in Group III, half of which ran to a black goal box on the runway and half of which ran to a white box, the goal boxes on the maze were either both black or both white, but different from the color box that had been used on the runway. The positive goal box was located on the nonpreferred side of the maze. For this group, as with Group I, the influence of previously established reinforcing value was eliminated in that the goal boxes on the maze were different from the goal box on the runway. For Groups I and III, then, the goal boxes on the maze possessed little, if any, reinforcing value at the beginning of the maze trials; for Group II, however, both the positive and the negative goal boxes already possessed some reinforcing value.

TABLE I

EXPERIMENTAL PROCEDURE

Groups	Color of Goal Box	
	Runway	Maze
I	grey	black-white
II	grey	grey-grey
III	black / white	white-white / black-black

TABLE II

MEANS AND *t*-RATIOS BETWEEN MEANS

Groups	Mean Number Correct Choices	*t*-Ratios	
		Group I	Group II
I	11.9	—	—
II	8.8	3.34	—
III	10.0	2.39	1.13

A non-correction technique was used on the maze. Fifteen trials were given each of the rats with an intertrial interval of approximately 10 min. When an incorrect response was made, the rat was retained in the goal box for 30 sec. and then returned to its home cage. Following correct responses the animals were given a cup of food and allowed to eat in the goal box for 30 sec. The experimental treatment for each group is summarized in Table I.

RESULTS AND DISCUSSION

The mean number of correct responses for each group of rats is shown in Table II. The rats in Group I that were run to dissimilar goal boxes on the maze have a mean of 11.9 correct responses. The rats in Group II that were run to the two similar goal boxes on the maze which were similar to the goal box used on the runway have a mean of 8.8 correct responses. The difference between these two group means is significant at the one percent level of confidence ($t = 3.34$). And the rats in Group III that were run to similar goal boxes on the maze which were different from the goal box used on the runway have a mean of 10.0 correct responses. The difference between this group and Group I is significant at the five percent level of confidence ($t = 2.39$), and the difference between this group and Group II is not statistically significant ($t = 1.13$).

In Fig. 3 learning curves are plotted showing the percent of correct responses as a function of the number of trials. Each point represents the

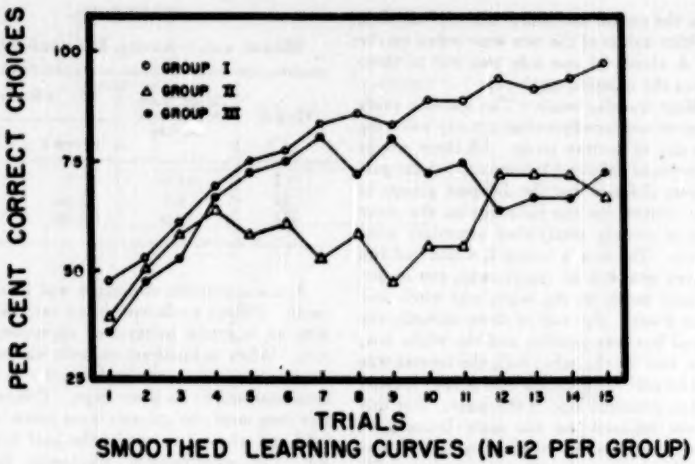


FIG. 3. Learning curves for all three groups of rats where each point is the average percent of correct responses for a given trial and the two adjacent trials

average percent of correct responses for a given trial and the two adjacent trials. It can be seen that the dissimilar goal box group (Group I) learned at a faster rate than either of the other two groups. Also, it can be seen that learning appears to be retarded earlier in the learning series for Group II than for Group III. A flattening of the curve occurs at trial four for Group II and at trial seven for Group III.

The results of this experiment support the generalization hypothesis which explains the more rapid learning with dissimilar goal boxes than with similar goal boxes. The learning is significantly slower when similar goal boxes are used than when dissimilar goal boxes are used, whether or not the negative goal box has already acquired reinforcing value before the learning proper is begun. The generalization of secondary reinforcement *established during the learning situation itself* is sufficient to produce the difference in learning speed between the rats learning the maze with similar goal boxes and the rats learning the maze with dissimilar goal boxes. However, when the acquisition of the secondary reinforcement does not occur until *after* the maze learning is begun, the hampering effect of the generalized reinforcing value in the maze situation is probably not as great as when the secondary reinforceing value is already established *before* the maze learning is begun. Although this statement is not supported by the experimental results in a statistical sense, the direction of the difference between the means of the two groups (Groups II and III), as well as the shapes of the learning curves, tends to support this notion. A small difference exists between the means of Groups II and III in favor of Group III.

SUMMARY

Three groups of 12 rats were given 25 reinforced runs on a closed runway over a period of five days. The rats in each group were then given 15 massed trials in a simple, closed

U-maze where choices of the nonpreferred side were rewarded. The two goal boxes on the maze were different for Group I; for Groups II and III they were alike. For Group I the two dissimilar goal boxes on the maze were both different from the goal box that had been used on the runway; for Group II the two similar goal boxes on the maze were both similar to the goal box that had been used on the runway; and for Group III the two similar goal boxes on the maze were both different from the goal box that had been used on the runway.

The results of the experiment support the generalization hypothesis which accounts for faster learning with dissimilar goal boxes than with similar goal boxes. When dissimilar goal boxes are used on the maze, learning proceeds at a faster rate than when similar goal boxes are used. This is true whether or not the experience previous to the maze learning leads to early acquisition of secondary reinforcement.

(Manuscript received May 23, 1949)

References

1. Denny, M. R. The effect of using differential end boxes in a simple T-maze learning situation. *J. exp. Psychol.*, 1948, 38, 245–249.
2. Saltzman, I. J. Maze learning in the absence of primary reinforcement: A study of secondary reinforcement. *J. comp. Psychol.*, 1949, 42, 161–173.

A PRELIMINARY DETERMINATION OF THE FUNCTIONAL RELATIONSHIP OF EFFECTIVE REACTION POTENTIAL ($_s\bar{E}_R$) TO THE ORDINAL NUMBER OF VINCENTIZED EXTINCTION REACTIONS ($\dot{n}$) [1]

BY HARDY C. WILCOXON, RUTH HAYS, AND CLARK L. HULL

Institute of Human Relations, Yale University

Introduction

Many of the functional relationships of reaction potential ($_sE_R$) to related phenomena were reported previously by Gladstone *et al.* (2). One such relationship not yet reported, however, is that between effective reaction potential and the ordinal number of extinction reactions. It is the primary purpose of the present article to present the first determination of this relationship, utilizing the method of quantification of reaction potential developed by Hull *et al.* (5). This method of quantification is reported in detail in the reference just cited, and will not be described here except to say that it consists of a modification and extension of Thurstone's paired-comparisons technique, Case III (6).

Strictly speaking, the relationship sought is that between *effective* reaction potential ($_sE_R$) and the ordinal number of Vincentized extinction reactions ($\dot{n}$). Perhaps a brief statement in definition of the symbols to be used is warranted. The symbol $_sE_R$ has been defined by Hull (4, p. 404) as "the potentiality of reaction evocation," or reaction potential. The symbol $_s\bar{E}_R$ is defined (4, p. 404) as *effective* reaction potential, i.e., reaction potential *minus* total inhibitory potential, the symbol for which is $\bar{I}_R$. Precisely, then,

$$_s\bar{E}_R = {}_sE_R - \bar{I}_R.$$

Now let us consider the symbol n. This has been defined (4, p. 406) as "the [total or cardinal] number of unreinforced reactions required to produce experimental extinction." In the present investigation, however, we are not primarily concerned with the total or *cardinal* number of reactions evoked in extinction. Instead, we are concerned with the *ordinal* numbers of the reactions evoked throughout the course of extinction. We shall therefore use a slightly modified symbol ($\dot{n}$) to designate these ordinal numbers.[2] In accordance with this symbolism, the relationship to be determined is,

$$_s\bar{E}_R = f(\dot{n}).$$

Subjects

The *S*s used in this study were 36 male hooded rats about 90 days old at the beginning of training. They were fed on a diet of dog biscuit supplemented weekly by an additional pellet soaked in codliver oil. The food used for reinforcement in the training consisted of soft round pellets

[1] This investigation was carried out in the laboratory of Professor Clark L. Hull as a part of the coordinated research program of the Institute of Human Relations. The study was done under the general direction of Professor Hull. Miss Hays performed all the experimental work and made the preliminary tabulations of the original data. Mr. Wilcoxon did the statistical work, the curve fitting, and wrote the manuscript.

[2] Specifically, $\dot{n} = 0$ will denote the reaction which has been preceded by no non-reinforced trials, but which results in the first non-reinforcement. Similarly, $\dot{n} = 1$ will designate the next reaction in the series, viz., the reaction which has been preceded by one non-reinforcement.

made from a mixture of two parts ground dog biscuit and one part flour, moistened with a little water. The animals were housed in round individual cages, 9 in. in diameter and 6.5 in. high.

Apparatus and Training Procedure

A diagram of the apparatus is presented in Fig. 1. It consisted of three compartments: an initial retaining compartment whose forward wall was a vertically sliding shutter; a small compartment leading from the shutter to a light, swinging door; and a large food compartment beyond the door. An electric timer was wired into the apparatus to record the time taken by the animal to go from the retaining compartment to the food compartment after the shutter was raised. The distance between these two compartments was eight in. The timer, which was calibrated in hundredths of a second, measured the interval between the instant the shutter was raised and the instant the swinging door leading to the food compartment began to open. The electric contact which started the timer was made when the shutter passed through a point one in. on its upward travel. This brief contact also activated a holding relay which kept the timer circuit closed until the animal pushed under the swinging door, where a micro-switch attached to the door broke the circuit, stopping the clock. The circuit was broken when the bottom of the door was moved outward one-half in.

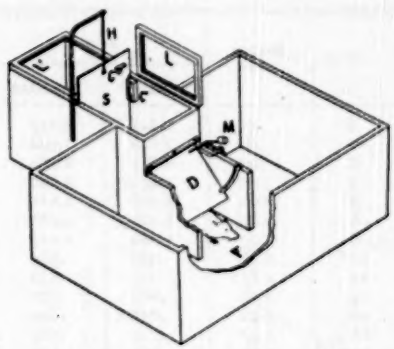


Fig. 1. The apparatus. At the beginning of each trial the rat was placed in the retaining compartment and the lid (L') closed. The experimenter then raised the shutter (S) and secured it on the hook at H, as shown in the figure. As the shutter moved upward, an electric timer (not shown) was started through the action of electrical contacts made by C at C'. When the animal pushed under the swinging door (D), the timer circuit was automatically broken by the microswitch at M. On the learning trials the rat always found a small pellet of moist food at F.

Preliminary training.—Throughout the investigation the animals were trained at 22 hours' food privation. The training for the first few days consisted of habituation of the animals to the apparatus and the feeding schedule. This involved the following procedure. An animal was first placed in the retaining compartment with a pellet of moist food like those used later in the training proper. After the animal had eaten his pellet the shutter was raised and he was allowed to explore the entire runway and the outside feeding compartment, the swinging door being propped open at this stage. There was never any food in the outside food compartment during the preliminary training. As soon as each rat ate promptly in the retaining compartment and moved about readily in the two other compartments, the swinging door was gradually lowered from trial to trial and the rat was trained to push under it and go out into the outside compartment. This training was continued until each animal would push open the completely lowered door and go into the outside compartment shortly after eating his pellet.

The training proper.—At this point the training proper began. The animal was placed in the starting compartment with the shutter down but without the customary food pellet. The shutter was then raised and the animal was allowed to walk through the middle compartment to the door, push under the door, and enter the large feeding compartment where he now found a food pellet. After he had eaten this pellet he was returned to his living cage for 20 to 30 sec. while the apparatus was arranged for the next trial and the latency of the completed trial was recorded. He was then given another trial. The animals were given 100 such trials.[3] The first 40 trials were given at the rate of four per day and the final 60 at the rate of 10 per day.

Extinction procedure.—Each animal was extinguished individually by being given massed trials without reinforcement. Since the extinc-

[3] The final 60 of these trials were given under slightly different conditions for three subgroups of the 36 animals. These differences involved having different stimulus cards attached to the response door (D, in Fig. 1), and were instituted for the purpose of investigating a problem not dealt with in the present article. Since *t*-tests of significance revealed that the different conditions did not differentially affect either the mean latencies or the median *n*'s of the three subgroups when they were subsequently extinguished, the three groups are here pooled.

tion of 36 animals necessarily took more than a single day to complete, each animal was given some retraining before his extinction trials began. On the third and second days before a given animal was extinguished he was given 4 retraining trials; on the day immediately preceding extinction he was given 10 retraining trials. On the day of extinction the rat was given one trial after another without any food pellets, until his latency of response exceeded five min. Latency was recorded on each trial.

Results

There was considerable variability in the number of trials required by the different animals to meet the 5-min. extinction criterion, the range being from 4 to 242 trials with the median at 44. In order to quantify the reaction potential by the present method, the different rates of extinction of the various animals had to be equalized. This equalization was accomplished by the justly weighted Vincent technique described by Hull (3, p. 245, footnote), which is a technique for changing a series of learning scores of any length to a comparable new series of any desired number of intervals without changing the mean of the series. Since the median number of trials to extinction for the entire group of animals was 44, the record of each of the animals was converted to a comparable Vincent series of 44 intervals. For these computations the criterion trial, i.e., the one on which the animal's latency exceeded five min., was not used.[4]

We did not think it feasible to scale the entire 44 intervals, so a selection of intervals from the series was made. The intervals selected for scaling are indicated by the dots in Fig. 2. The general plan governing the selection was to take alternate intervals except in the regions near the beginning and near the end of extinction where change could be expected to be more rapid and consequently the particular intervals involved most important. This selection yielded a total of 26 intervals which were then scaled by the method cited above. Since previous experience (7) with the use of this scaling method had demonstrated that data of this kind are subject to a distortion when the normal distribution table is used in their quantification, a special table for leptokurtic distributions was used. The actual

[4] The use of the criterion trial in the Vincentizing procedure would have inflated the last few Vincent latencies ($_{sL}n$'s), especially for those animals whose n was less than 44.

TABLE I

Ordinal Number of Vincentized Extinction Intervals (n), the Corresponding Vincentized Reaction Latency Medians, the Parallel Effective Reaction Potentials as Calculated by a special Method Described in Detail Elsewhere (5), and the Theoretical Effective Reaction Potentials Calculated from the Equation, $_{s}\bar{E}_{R} = 1.84 \times 10^{-.021\,6} + .277$

n	Median ($_{sL}n$)	$_{s}\bar{E}_{R}$	
		Empirical	Calculated
0	.52	2.145	2.117
1	.60	2.023	1.942
2	.92	1.824	1.784
3	3.07	1.442	1.640
4	3.61	1.307	1.511
6	4.04	1.303	1.287
8	4.45	1.049	1.104
10	5.26	.929	.954
12	4.85	.853	.832
14	4.60	.806	.731
16	5.42	.831	.649
18	5.47	.703	.582
20	6.64	.564	.526
22	10.91	.387	.481
24	9.24	.434	.444
26	9.98	.291	.414
28	8.09	.436	.389
30	5.45	.628	.369
32	9.20	.287	.352
34	7.28	.518	.339
36	10.47	.330	.327
38	9.78	.290	.318
40	9.10	.448	.311
41	10.48	.444	.308
42	19.86	−.020	.305
43	23.58	.000	.302

table was identical to the one published by Yamaguchi *et al.* (7, p. 225).

After the data were scaled, a calculation of the distortion still remaining in the computations in spite of the use of this leptokurtic table, revealed that the amount of distortion was no more than could be expected by chance 21 times out of a hundred ($t = 1.38$, $df = 23$). The tendency toward distortion which did exist, however, was in the direction of indicating that the data quantified were more leptokurtic than the table (cf. Yamaguchi *et al.*).

The resulting values of ${}_s\bar{E}_R$ at the various Vincentized intervals throughout the course of extinction are shown in Table I. In order to obtain a mathematical expression of the relationship being sought, a curve was fitted to these values. By following the procedure of successive approximations described by Felsinger *et al.* (1, pp. 224–226), the best-fitting curve found was,

$$ {}_s\bar{E}_R = 1.84 \times 10^{-.0434\,\dot{n}} + .277. $$

This equation is shown plotted as a smooth line drawn through the scaled ${}_s\bar{E}_R$ values in Fig. 2. The theoretical ${}_s\bar{E}_R$ values derived from this equation, and corresponding to the intervals which were scaled, are presented in Table I. The mean square deviation of the theoretical from the empirical values is .019 σ.

Discussion

It will be noticed that the last two dots in Fig. 2, i.e., those representing the scale values at $\dot{n} = 42$ and $\dot{n} = 43$, do not appear to be consistent with the *gradual* trend downward of the other dots taken as a whole. The decrement in ${}_s\bar{E}_R$ at these last two intervals seems to be rather more precipitous. This same phenomenon is revealed also in the relatively abrupt increase in the median Vincentized latencies corresponding to these intervals (cf. Table I). At first glance it might appear that this deviation from the general trend of the data is the result of some artifact produced by

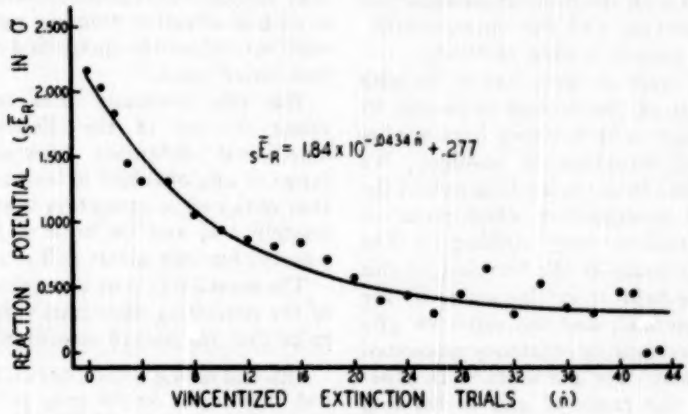


Fig. 2. Graph showing the functional relationship between effective reaction potential (${}_s\bar{E}_R$) and the ordinal number of Vincentized extinction reactions ($\dot{n}$). The dots represent the scale values of ${}_s\bar{E}_R$ at 26 Vincent intervals in the extinction process, as derived by our present method of quantification (Hull *et al.*, 5). The smooth curve drawn through these data represents the equation fitted to the empirical values.

the treatment of the data; perhaps by the Vincentizing. This possibility would seem particularly likely had the 5-min. criterion trial been retained in each of the 36 individual series of scores which were fractionated (cf. footnote 4). For in that event, the normal amount of variability in latency from one animal to another at any given trial would have been restricted drastically at the last interval. This would have created the inverse of what is frequently seen as an end spurt in Vincentized learning curves. However, these considerations are ruled out as an explanation in view of the fact that the criterion trial was not used in calculating the Vincent series. The exact explanation of the phenomenon is therefore unknown.

An interesting characteristic of the equation fitted to these data is its relatively narrow range in ${}_{S}\bar{E}_{R}$ value. The total range covered by the equation, i.e., from its highest value at $n = 0$ to its asymptotic value of .277, is only 1.84 σ units. This fact raises several interesting problems in connection with the possible generality of the function and the characteristics of the present scaling method.

The main problem has to do with the unit of the derived scale and its constancy or inconstancy from one behavioral situation to another. We have data from the learning part of the present investigation which point up this problem very strikingly. The first 40 trials in the learning of this running habit were also scaled by the same method, and the value for ${}_{S}E_{R}$ at the asymptote of latency amounted to slightly over 4 σ units.[5] In other words, the range of ${}_{S}E_{R}$ in learning was from zero to approximately 4 σ, or a range of 4 units, whereas the range in extinction was only 1.84 units. Unfortunately these animals were not extinguished immediately after the original learning, but were run for 60 additional trials under slightly changed conditions as previously noted (cf. footnote 3). During this intervening period, the median latency of the group increased from 0.34 sec. on the fortieth trial to 0.46 sec. on the first extinction trial.[6,7]

The question of immediate interest, therefore, is: how much of the discrepancy in range between our learning ${}_{S}E_{R}$ and extinction ${}_{S}\bar{E}_{R}$ can be accounted for by this difference in latency?

An estimate may be obtained by considering the first trial of extinction as theoretically equivalent, as far as latency is concerned, to an additional learning trial. When this was done and the median latency for the first extinction trial was substituted in the equation ${}_{S}E_{R} = f({}_{S}t_{R})$, which we fitted to the learning data, a value of 3.25 σ was obtained. This means we may estimate that approximately one σ unit of effective reaction potential was lost before the quantified extinction series began.

But this obviously does not account for all of the discrepancy. The total difference between the range of ${}_{S}E_{R}$ obtained in learning and that obtained in extinction is approximately 2 σ, and we have so far accounted for only about half of it.

The most likely possible explanation of the remaining discrepancy appears to be that the unit of measurement in

[5] A detailed presentation of this part of the investigation will not be made here, since the scaling of a somewhat similar learning process has already been published (2).

[6] The trial referred to is the first non-Vincent trial. The latency for the group at the first Vincent interval was 0.52 (cf. Table I).

[7] This gradual increase in latency and the parallel decrease in reaction potential resulting from continued reinforced trials beyond the asymptote of learning were observed also by Yamaguchi *et al.* (7).

the learning scale is different from that in the extinction scale. Since the unit of these scales is based on the variability in performance of the group of organisms used as subjects (5), there is excellent reason to expect that the unit itself will vary in magnitude between situations in which there is different variability among the subjects. It is quite possible, therefore, that this unit could vary from one situation to another and especially from one type of behavioral process (learning) to another (extinction).[8] It must be pointed out, however, that conclusive evidence bearing upon this fundamental question awaits research which is specifically directed to an investigation of the problem. Nevertheless, the present evidence is sufficient to reveal the tentative nature of our present equation, particularly in regard to the coefficient (1.84 σ).

Summary

1. The first determination of the functional relationship between effective reaction potential ($_{s}\bar{E}_{R}$) and the ordinal number of Vincentized extinction trials (n) is presented, the method of quantifying reaction potential developed by Hull *et al.* (5) having been utilized. The mathematical expression of this relationship is,

$$_{s}\bar{E}_{R} = 1.84 \times 10^{-.0434\,n} + .277.$$

2. This preliminary determination is based on a paired-comparisons quantification of the latency data of 36 male hooded rats which were given massed extinction trials on a simple running response.

3. Evidence obtained from a similar quantification of the original learning of the running response indicates that the unit of the $_{s}E_{R}$ scale as a whole probably varies in magnitude from learning to extinction.

4. It is pointed out that this evidence renders the present determination of $_{s}\bar{E}_{R} = f(n)$ quite tentative in respect to the coefficient of the right-hand member of the equation.

(Manuscript received May 19, 1949)

[8] The σ_4 procedure employed in Thurstone's Case III, which we followed, presumably equalizes the dispersions at the several points scaled *within* each series of items scaled. It does not, however, insure the equality of dispersions (and, therefore, of the scale unit) from one scaled series to another.

References

1. Felsinger, J. M., Gladstone, A. I., Yamaguchi, H. G., & Hull, C. L. Reaction latency ($_{s}t_{R}$) as a function of the number of reinforcements (N). *J. exp. Psychol.*, 1947, 37, 214–228.
2. Gladstone, A. I., Yamaguchi, H. G., Hull, C. L., & Felsinger, J. M. Some functional relationships of reaction potential ($_{s}E_{R}$) and related phenomena. *J. exp. Psychol.*, 1947, 37, 510–526.
3. Hull, C. L. Simple trial-and-error learning —an empirical investigation. *J. comp. Psychol.*, 1939, 27, 233–258.
4. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century Co., 1943.
5. Hull, C. L., Felsinger, J. M., Gladstone, A. I., & Yamaguchi, H. G. A proposed quantification of habit strength. *Psychol. Rev.*, 1947, 54, 237–254.
6. Thurstone, L. L. Stimulus dispersion in the method of constant stimuli. *J. exp. Psychol.*, 1932, 15, 284–289.
7. Yamaguchi, H. G., Hull, C. L., Felsinger, J. M., & Gladstone, A. I. Characteristics of dispersions based on the pooled momentary reaction potentials ($_{s}\bar{E}_{R}$) of a group. *Psychol. Rev.*, 1948, 55, 216–238.

EFFECTS OF COMPETING REACTIONS ON THE CONDITIONING CURVE FOR BAR PRESSING*

BY WILLIAM K. ESTES

Indiana University

Introduction

In his original studies of bar-pressing behavior in the rat, Skinner (4) noted appreciable variability among animals with respect to form of the original acquisition curve. Under regular reinforcement, some animals achieved a maximal rate of responding almost instantaneously, while others reached final constant rates only after a considerable period of initial acceleration. Similar variability has been found in other conditioning situations (see, for example, 3, 145–148). A quantitative formulation of conditioning, which has been developed elsewhere (1), implies that a principal source of variation in form of acquisition curves is the concurrent extinction of competing reactions. Direct verification of this assumption is limited by the virtual impossibility of identifying and measuring all possible competing reactions in any learning situation. Indirect evidence may be obtained, however, by the following type of experiment.

Suppose that for some response in a standard experimental situation we are able to find conditions under which conditioning curves without marked initial acceleration may be regularly obtained. Then by arranging conditions so that a strong competing reaction will be extinguished while the other response is being reinforced, we may determine whether the resulting curve is similar in quantitative properties to ordinary acquisition curves with initial acceleration. This approach is illustrated in Fig. 1, where we have plotted mean cumulative curves of bar-pressing respones vs. time for a group of three rats run in a Skinner box. Motivation was 24 hour thirst, and all responses were reinforced with water. Curve R_1 represents the first period, during which the rats were conditioned to press a bar on the right side of the box, when only that bar was present. The curve shows the initial acceleration which has been frequently obtained during original conditioning of bar-pressing responses in this apparatus. These animals were given 100 reinforcements per daily period for five days, the mean curve for the fifth day being R_5 in the figure. It will be noted that the curve on the last day is virtually linear, with a relatively high and constant rate of responding maintained almost from the beginning of the

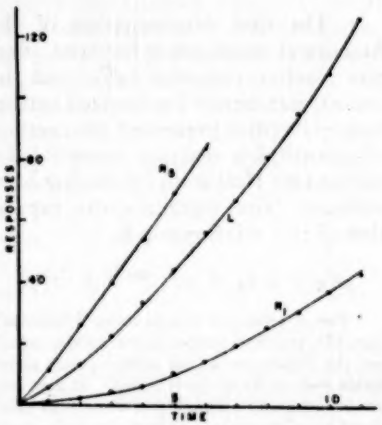


Fig. 1. Mean cumulative records of bar pressing for three rats. Curve R_1 represents original conditioning to the right bar; curve R_5, responses to the right bar during the fifth period of reinforcement; curve L, subsequent conditioning to the left bar. The smooth curves are freehand graduations.

*This study was supported in part by a research grant from the Graduate School of Indiana University.

period. On the sixth day, the right bar was removed and a similar bar inserted on the left side of the box, depressions of this bar now being reinforced with water. The mean record is represented by curve L in the figure. As might be expected, conditioning was much more rapid than original conditioning on the first bar, a relatively brief period of initial acceleration reaching a rate nearly equal to that attained on the right bar after five days of training. Now, if we are correct in assuming that the amount of initial acceleration is determined primarily by the relative initial strengths of the reinforced response and other competing responses, the introduction of a strong competing reaction on the sixth day should result in an acquisition curve for the left bar very similar to the curve of original conditioning on the right.

In the experiments to be described, rats are conditioned to press a bar for water reinforcement. Then they are permitted to learn a new bar response while the first habit is being extinguished. The curves of conditioning obtained for the two bar responses will be examined for any differences in form which may be produced by the differences in experimental conditions. Similarity of the two acquisition curves will be taken as evidence for the view that acceleration in the first curve was determined, at least in part, by concurrent extinction of unidentified competing responses.

In order to make exact quantitative predictions, we shall use the following equation, which has been shown to describe changes in rate of occurrence of an operant response under regular reinforcement (for derivation of the function, see (1)):

$$(1)\quad r = \frac{w}{1 + \frac{(w - r_0)}{r_0} e^{-Bt}}$$

where r represents rate of responding at time t. The slope constant, B, is the product of w and a constant q which may be thought of as the proportion of stimuli in the experimental situation effective at any given moment. The rate of responding will increase from its initial value, r_0, and eventually approach a maximum value, w. The presence or absence of positive acceleration in the early portion of the curve will be determined by the relative magnitudes of r_0 and w. Other conditions being constant, one important way in which the relative magnitude of r_0 may be modified is the strengthening or weakening of competing responses. In the following experiment, if the assumptions underlying this formulation have been approximately realized by the experimental conditions, we may predict that the acquisition curves of the first and second bar responses will differ only with respect to the magnitude of r_0 in equation (1). To apply this equation directly to our experimental data, we should have to estimate rates of responding after various intervals of time from the obtained records. This could be done, but would involve an unnecessary loss of accuracy. Instead, we can integrate equation (1) to obtain an expression for total number of responses, R, as a function of time in minutes, t, since the beginning of learning, thus:

$$(2)\quad R = wt + \frac{w}{B} \log \left(\frac{r_0}{w} + \frac{(w - r_0)}{w} e^{-Bt} \right)$$

The integrated form of the equation will be utilized in analyzing the data of the present study.

Procedure

The apparatus, consisting of four Skinner-type conditioning units, has been described in detail elsewhere (2). Each rat was enclosed during the experimental period in a partially sound-proofed chamber containing a magazine which could be set to deliver 5 mg. of water. Two horizontal aluminum bars could be made

available to the animal through slots in the wall of the experimental cage, one on either side of the magazine. Each bar was 1¾ in. in length, and projected ¾ in. from the wall at a height of 4 in. Depression of either bar a distance of $\frac{1}{16}$ in. with a force of 4½ gm. actuated a rotary microswitch, permitting electrical impulses to go to the recorder and, when the procedure called for reinforcement, to the magazine motor. Whenever either bar was present in the animal's compartment, a cumulate curve representing depressions of that bar vs. time was recorded. The experimental cage could be removed from the soundproof chamber, and served as the animal's living cage between experimental periods, being housed in a constant temperature cabinet maintained at 26°C.

Subjects were 11 male albino rats obtained from an unselected stock maintained by a local commercial vendor. The four rats used in Experiment 1 were 110 days old at the beginning of the experiment. (The three rats whose records were cited in the Introduction, above, were of the same age and were run concurrently with the subjects of Experiment 1; results obtained with these three animals will not be discussed further in the present report.) The four rats used in Experiment 2 were 180 days old. Dry food was available to the animals at all times; water was available in the living cages for one hour immediately after each daily experimental period.

Experiment 1.—The subjects were first habituated to the situation and trained to respond to the sound of the magazine by approaching and drinking the water, 150 'free' reinforcements being distributed over eight days; the bars were not present in the boxes during these periods. Then both bars were inserted for a 30-min. period during which no reinforcements were given, in order to extinguish most of the unconditioned exploratory behavior which might otherwise contribute undesired variability to the conditioning curves. On the next day, the right bar only was inserted and 100 bar-pressing responses were reinforced with water; 100 additional reinforcements were given for responses to this bar on each of the next four days. On the last day of the experiment, both bars were present, all responses to the left bar being reinforced while responses to the right bar were extinguished for a 30-min. period. Responses to the two bars were recorded separately.

Experiment 2.—This consisted in a replication of the first experiment under slightly modified conditions. The subjects were first habituated to the apparatus and trained to drink from the magazine, receiving 110 'free' reinforcements distributed over nine daily periods. Then in order to collect data for another purpose, these rats were satiated on both food and water and placed in the apparatus with the left bars present for 30 min., bar-pressing responses being followed by the sound of the empty magazine; it might be noted that very few responses were obtained, the mean being only 9.5. On the next day, the left bars were again present and all bar-pressing responses were reinforced with water, the period continuing until the animals were satiated; this required about four hours, some of the curves containing over 2000 responses; on the day following, the animals were not run but were given water for one hour at the usual time. On the final day, both bars were present in the boxes, all responses to the right bar being reinforced while responses to the left bar were extinguished.

Results

Only acquisition data will be considered in this paper. (The extinction curves obtained for the first-learned response during acquisition of the second will be reported elsewhere). In order to put the data in suitable form for curve-fitting, the cumulative kymograph records for the animals of each group under each condition of acquisition were measured at the end of each minute following the

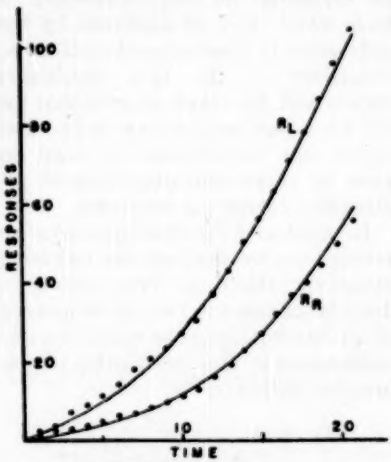


Fig. 2. Mean cumulative records of bar pressing during original conditioning to the right bar, R_R, and subsequent conditioning to the left bar, R_L, of Experiment 1. The smooth curves represent theoretical equations given in the text.

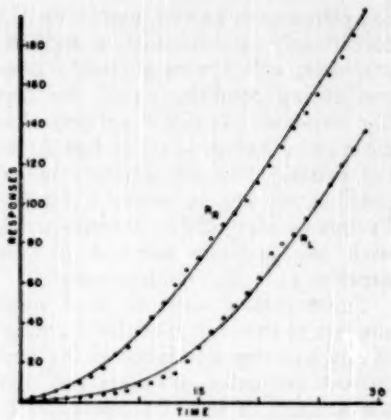


FIG. 3. Mean cumulative records of bar pressing during original conditioning to the left bar, R_L, and subsequent conditioning to the right bar, R_R, of Experiment 2. The smooth curves represent theoretical equations given in the text.

first response, and averaged. These mean values of cumulative responses vs. time are plotted as the empirical points in Figs. 2 and 3.[1] It should be emphasized again that the curves for the second conditioning in each experiment *do not include* unreinforced responses to the other bar. Equation (2) has been fitted to the original acquisition curves of the right bar habit in Experiment 1 and the left bar habit in Experiment 2, and the theoretical curves included in Figs. 2 and 3 respectively. Evaluation of the constants by a method of approximations yielded the following values:

$$\text{Exp. 1} \quad R_R = 9.6t + 122.7 \log_{10}(.045 + .955e^{-.18t}) + 1$$

$$\text{Exp. 2} \quad R_L = 9.4t + 86.5 \log_{10}(.026 + .974e^{-.25t}) + 1$$

[1] For a table of these data and values computed from the equations mentioned below order Document 2740 from American Documentation Institute, 1719 N Street, N.W., Washington 6, D. C., remitting $0.50 for microfilm (images one in. high on standard 35 mm. motion picture film) or $0.50 for photocopies (6 × 8 in.) readable without optical aid.

One response has been added to each equation to allow for the fact that in obtaining average curves the individual records were measured after intervals of time counted from the first response rather than from the beginning of the experimental period.

Inspection of the figures reveals that for Exp. 1 the equation fits the data adequately for the first 17 min., after which the empirical points fall appreciably below the curve; this systematic deviation is caused by an abrupt decline in rate on the part of one rat, and on the assumption that it was due to the effect of a change in motivation, has not been allowed to modify the values of the constants obtained in fitting the first part of the curve. In the case of Exp. 2, the theoretical curve seems to provide a fair graduation of the observed values.

In fitting the conditioning data of the second bar response in each experiment, the value of r_0 has been estimated from the data, while the values of the other constants have been taken from the equations fitted to the acquisition curves of the first bar response, yielding

$$\text{Exp. 1} \quad R_L = 9.6t + 122.7 \log_{10}(.13 + .87e^{-.18t}) + 1$$

$$\text{Exp. 2} \quad R_R = 9.4t + 86.5 \log_{10}(.13 + .87e^{-.25t}) + 1$$

These theoretical curves have been plotted in the figures.

In order to provide some quantitative expression of the goodness of fit of the various curves, differences between theoretical and observed values of R have been computed. The standard deviations of these residuals (standard errors of estimate) for the first and second conditionings of Exp. 1 are 1.23 and 1.17; and for Exp. 2, 2.45 and 2.44 respectively. The squares of these standard deviations, together with the variances of the

observed values of R, may be entered in the formula $100\left(1 - \frac{\sigma^2_{R,t}}{\sigma^2_R}\right)$ to obtain for each record an estimate of the percentage of the variance of the observed R values which is accounted for by the theoretical curve. The values obtained, in the same order as the standard errors just listed, are 99.43, 99.86, 99.69, and 99.84 percent. Although these statistics should be interpreted with some caution, since we are not prepared to demonstrate that all assumptions underlying the use of the index of determination have been fully met, it seems safe to conclude that the goodness of fit of the theoretical equations is at least as good for the second acquisition curve as for the first in each experiment.

Discussion

Although not essential to the interpretation of this experiment, a conjecture as to the specific nature of the competing responses which are believed to retard acceleration of original acquisition curves may be of some interest. In experiments involving bar pressing and food or water reinforcement, it is necessary to train the animal to respond to the magazine before bar responses can be conditioned. It seems likely that during this preliminary training various responses may be 'fortuitously' reinforced and develop considerable strength. Skinner (5) has demonstrated experimentally that elaborate 'superstitious' behaviors may develop under these conditions. When, subsequently, reinforcement is made contingent upon bar pressing, the other responses may be expected to extinguish while the rate of bar pressing increases. In the second conditioning periods of Experiments 1 and 2, above, we have substituted a competing response of known history for the fortuitously conditioned competing responses which were probably present during conditioning of the first bar response. It is not unlikely that close examination of other conditioning situations would similarly reveal possible sources of strength for behaviors which would be in competition with the response selected by the experimenter for reinforcement.

These results seem to lend some support to the view that the 'learning' of any one response involves the concurrent extinction of others and that the amount of initial acceleration in a learning curve is determined to an important extent by the relative initial strengths of all behaviors which may occur in the experimental situation.

If there are any differences in form of the acquisition curves obtained for bar pressing responses with and without concurrent extinction of another strong response, the differences are too small to be revealed by our quantitative analysis of the data. On the assumption that any marked variation in experimental conditions should be reflected in the quantitative properties of the conditioning curve, it is suggested that the first and second learning situations in each of the experiments described above differed only in that during the original learning the extinction of competing behavior was not recorded by the experimental apparatus.

Relationships between the first and second curves of conditioning for each group tend to verify a number of exact quantitative predictions which follow from a statistical theory of simple learning. The acquisition curve for bar pressing derived from this theory involves three parameters: w, the reciprocal of the minimum time required to execute the specified re-

sponse; q, reflecting a relationship between properties of the stimulating situation and the organism's receptive apparatus which should be relatively constant throughout a learning period; and r_0, the initial rate of the response to be learned. As predicted on the basis of this formulation, only r_0 changed in value from the first to the second learning curve within each experiment; thus it was possible to predict with considerable accuracy the course of acquisition of a second bar response, knowing only its initial rate and the parameters of the conditioning curve of a previously learned response.

Summary

These experiments were designed to obtain evidence as to the role of competing reactions in modifying the form of curves of conditioning. The apparatus consisted of four Skinner boxes each equipped with a magazine and two bars differing only with respect to position in the box. Two groups of rats were first conditioned to press one bar for water reinforcement, then conditioned to press the second bar while the first learned response was being extinguished. Acquisition curves obtained under the two conditions appear virtually identical in form, differing only with respect to initial rates of responding. In each experiment, the course of the second conditioning could be predicted accurately from a knowledge of the initial strength of the second response and the parameters of the first conditioning curve. This evidence is regarded as supporting the view that in general the form of conditioning curves is determined to an important extent by the concurrent extinction of competing reactions.

(Manuscript received April 11, 1949)

References

1. Estes, W. K. Toward a statistical theory of learning. *Psychol. Rev.* (In press)
2. Guttman, N., & Estes, W. K. A modified apparatus for the study of operant behavior in the rat. *J. gen. Psychol.* (In press)
3. Hilgard, E. R., & Marquis, D. G. *Conditioning and learning.* New York: Appleton-Century, 1940.
4. Skinner, B. F. *The behavior of organisms.* New York: Appleton-Century, 1938.
5. Skinner, B. F. 'Superstition' in the pigeon. *J. exp. Psychol.*, 1948, 38, 168–172.

THE EFFECT OF PERCEPTION TIME UPON RIGIDITY AND CONCRETENESS OF THINKING

BY MILTON ROKEACH

Michigan State College

Introduction

The present investigation arose out of an earlier study (22) of the relationships existing among prejudice, rigidity and concrete thinking. In that study it was demonstrated that high-prejudiced adults and children, as determined by their responses to a prejudice scale, were significantly more rigid in solving various non-social problems than low-prejudiced adults and children. It was demonstrated, further, that those scoring above the median in prejudice were not only more rigid in their problem-solving behavior, but also that the mode of thought leading to the solution was significantly more concrete in their case than in the case of subjects falling below the median on the prejudice scale.

Rigidity was defined as "the inability to change one's set when the objective conditions demand it, as the inability to restructure a field in which there are alternative solutions to a problem in order to solve that problem more efficiently" (22, p. 260).

Concreteness of thinking was defined operationally in terms of (1) the frequency of the use of scratch paper as an aid in solving certain arithmetic problems which involved the subtraction of one- and two-digit numbers from other two-digit numbers and (2) in terms of the number of words used to write down the solutions to both arithmetical and 'map' problems.

The results of the study just referred to raised the following question for further investigation: What is the nature of the personality factors which will help account for the differences in rigidity and concrete thinking found between those high and those low in prejudice?

One feasible way to attack this problem is to shift the focus of attention from the study of personality differences between the two groups to the study of the determinants of rigidity itself. It was anticipated that such an approach would serve a three-fold purpose: First, it should yield information about the nature of certain structural determinants of rigidity as such. Second, it should be possible from such an analysis to infer the existence of certain personality differences between rigid and non-rigid individuals—a problem important for both personality theory and clinical psychology. Third, to whatever extent rigidity and prejudice are related, to that extent it should also be possible to infer the existence of certain personality differences between high- and low-prejudiced individuals.

If a person is confronted with a problem, social or non-social, which can be solved either rigidly or flexibly, he will behave in one way or the other depending upon both structural and functional determinants of his perception of the problem situation. Such factors as brain-injury (4, 9, 23, 24, 30), decortication (2, 3, 11, 13, 17), feeblemindedness (12, 15, 28), deafness and blindness (18) would be considered examples of primarily structural determinants. Another such determinant, non-pathological in nature, is repetition. Thus, Krechevsky and Honzik (14) have shown with rats and Luchins (16) has shown with humans that fixation or rigidity in-

creases with the repetition of a given pattern of behavior.

More relevant to the present inquiry is the availability of concrete or abstract 'tools' as structural determinants. Disregarding for the moment the question of individual differences in the use of concrete 'tools' it may well be supposed that the individual having available only concrete 'tools' will behave more rigidly, other things being equal, than another individual having available more abstract 'tools.' Thus, in the writer's earlier study, it was evident that a subject using scratch paper as a concrete aid was more frequently rigid in his solution of relatively simple arithmetic problems than another subject not utilizing such a concrete aid.

The studies by Goldstein and Sheerer (10), Lewin (15), Werner (29), and Vigotsky (27) may be cited here in support of the hypothesis that concreteness of thinking is one important structural pre-condition of rigid behavior.

This raises a further problem: What determines concrete-abstract thinking? In the present investigation attention will be directed primarily toward one variable which is hypothesized to be an important structural determiner of concrete-abstract thinking, namely, the time element. It is apparent that abstract thinking can take place only when there is time available to remove oneself sufficiently from the immediate problem situation. If, for some reason, a person cannot generally tolerate delays or happens to be in a hurry, he will simply not have the time that is necessary to think through a problem on an abstract level.

The hypothesized importance of time availability for abstract thinking, however, is only part of the story. Time availability is probably also important for perception. Perhaps a more adequate formulation, which will include what happens in perception as well as in thinking, is this: The person in a hurry will be forced to *perceive* a problem confronting him narrowly and, simultaneously, will be forced to resort to behavioral supports of a concrete nature in order to solve the problem. Conversely, the person in no hurry at all will have time to *perceive* the problem more broadly and, simultaneously, will solve the problem more abstractly. Perception and thinking are regarded as being so dynamically interrelated that it is frequently difficult to say exactly where perception ends and thinking begins. Perhaps thinking can be no better and no worse than the perception on which it is based (and vice versa).

One consequence of having or not having enough time for perceiving-thinking is flexibility or rigidity of behavior.

The purpose of the present experiment is to test the hypothesis just presented but it is first necessary to consider the results of certain experiments reported by Luchins (16). Luchins gave *Ss* a series of arithmetic problems to solve under normal and under speed conditions. Under normal conditions the *Ss* were given 2½ min. per problem, which was considered an adequate amount of time for most *Ss*. Under speed conditions the *Ss* were given considerably less time per problem. The results showed that the groups solving the problems under speed conditions were considerably more rigid than the groups solving the problems under normal conditions. Such results are in accordance with the hypothesis presented here with respect to the effect of (lack of) time availability upon rigidity but provides no evidence as such about such effects upon concreteness of thinking.

If a decrease of time availability (as is the case under speed conditions) effects an increase in rigidity, then, according to the present hypothesis, it should be expected that an increase in time availability will effect a decrease in rigidity. But the existing evidence seems to indicate, at first glace, that such is not the case. In yet other experiments Luchins (16) eliminated completely the speed factor by explicitly instructing his subjects that there was no need to hurry and by giving his subjects plenty of time (3 to 10 min.). Compared with groups solving the problems under normal condi-

tions (2½ min.) there appeared to be no significant decreases in rigidity in the non-speed groups.

The results of these non-speed experiments, if confirmed, would run counter to the present hypothesis concerning the importance of time availability as a structural determinant of concrete thinking and rigidity, as a functional determinant of differences between rigid and non-rigid individuals and of possibly similar functional differences between high-prejudiced and low-prejudiced individuals. It is, therefore, necessary to consider in somewhat greater detail what probably happened on the cognitive level in these experiments.

In Luchins' speed experiments, it may be suggested, the speed factor served the function of preventing broad cognitions of the problem to a generally greater extent than under the normal conditions. Consequently, the subjects under the former conditions solved the problems more rigidly than under the latter conditions.

In Luchins' non-speed experiments, however, although the *S*s were *given* more time by *E* to solve the problems, no experimental steps were taken to ensure that they actually *'took'* this time to perceive the problem. From the phenomenological point of view, it may be suggested that the *S*s probably had no more time available to perceive under the non-speed conditions (3 to 10 min. per problem) than under normal conditions (2½ min. per problem). Under the non-speed conditions the *S*s 'took' the extra time to engage in post-solution activities like checking computations, writing solutions more neatly, etc. rather than in the more crucial pre-solution activity of *perceiving* the problems. In other words, it is the writer's contention that in spite of the great differences in total time given by *E* under the two sets of conditions the length of time spent in perceiving activity was approximately the same. Under both sets of conditions each *S* 'took' whatever time his personality structure normally 'permitted' him to 'take' and all additional time was phenomenologically superfluous. Consequently, no significant differences in rigidity were evident between the two conditions.

The above theoretical considerations led to an attempt to design an experiment in which the effect of systematic variation of time availability to perceive could be studied. It was predicted that increases in time availability to perceive should (1) effect decreases in rigidity and (2) effect decreases in concrete-thinking.

Procedure

The measurement of rigidity.—The technique used was one utilized by Luchins (16) in his study of the Einstellung-effect and also by the writer (22). The *S*'s task is to solve different problems in which required quantities of water are to be obtained by manipulating three jars of given capacities. A set is established by first presenting several problems which can be solved only by one relatively complicated technique. These are followed by critical problems which can be solved not only by the complicated method but also by a more direct method. An example of the first type of problem, capable of only one complicated solution, is as follows: Given are three jars with capacities of 31 quarts, 61 quarts and 4 quarts, respectively. The problem is to get 22 quarts of water using these three jars. If the three jars are labelled A, B, and C, the solution is B − A − 2C (61 − 31 − 4 − 4 = 22).

The above set problems are followed by several critical problems solvable not only by the B − A − 2C method but also by a more simple method. For example, get 20 quarts with jars that hold 23, 49, and 3 quarts. It will be seen that this problem is capable of two solutions: (1) B − A − 2C (49 − 23 − 3 − 3 = 20) and (2) A − C (23 − 3 = 20). An *S* who gives the B − A − 2C solution to a critical problem, in which a simpler solution is possible, may be said to be solving such a problem in a rigid manner.

The measurement of concrete-thinking.—As in the writer's earlier study two measures of concrete-thinking were available: (1) The *S*s were given the opportunity, if they so desired, to use scratch paper for figuring their solutions to the various problems. It is thus possible to compare different experimental groups on the frequency of use of scratch paper as a concrete aid in solving each problem. (2) The *S*s were also given the opportunity, if they so desired, to write out their solutions either in arithmetical

form or in sentence form. Differences in the mean number of words used by each experimental group may be tken as a further indication of differences in concrete-thinking.

Instructions.—"This experiment is designed to see how well you can work simple problems without making mistakes. Your performance on this experiment will in no way affect your grade in this course.

"Imagine that you are standing beside a lake which has lots of water. In the first problem you are given three empty jars. The first empty jar has a capacity of 13 quarts of water. The second empty jar has a capacity of 29 quarts of water. The third empty jar has a capacity of 3 quarts of water. Write down in your own words on the first page of your blue-book how you would go about getting exactly 10 quarts of water. Do the problem in the easiest possible way. You do not have to use all of the three jars in order to solve the problem. [Allow 2 min., then present Problem 2, allowing 2 min.]

"I will now explain how to do this problem [Problem 2]. The correct solution is to fill the 37-quart jar and from this to fill the 17-quart jar once and the 6-quart jar twice. There then remains in the 37-quart jar the required 8 quarts of water.

"Here is another way the solution to this problem may be written: $37 - 17 - 6 - 6 = 8$.

"Each problem which I will give you is to be worked on a new right-hand page of your blue-book. You will be allowed exactly 2 min. to solve each problem. You are not allowed to turn any of the pages in your blue-book except when I tell you to. Also, you are not allowed to go back to any previous problem on a previous page to change your answer or to fill in an omitted answer.

"There is one more important rule which you must follow. You are not allowed to start writing down your solution until I say 'Go.' Remember, this is important. Wait until I say 'Go' before you start to write down your solution to the problem. While you are waiting for my signal look at the problem on the blackboard.

"Turn over to the next right-hand page. There will be 9 more problems for you to solve. You may use the left-hand side of your blue-book for figuring and scratch work. Here is the next problem, Problem 3."

Experimental groups and design of problems used.—Four experimental groups were used—a 10-sec. Delay Group, a 20-sec. Delay Group, a 30-sec. Delay Group and a 60-sec. Delay Group. The *S*s, tested in groups of from 40 to 60, were taking courses in elementary and applied psychology at Michigan State College. They were assigned randomly, as classroom groups, to one of the four experimental groups. The differences in experimental conditions for these groups and the design of the problems used are shown in Table I.

The problems were presented one at a time on the blackboard. Problem 1 is a *control critical* problem solvable both by the complex and simple methods. Those few subjects in each experimental group who solved this problem by the complex (rigid) method were elimi-

TABLE I

DESIGN OF PROBLEMS USED AND DIFFERENCES IN TIME CONDITIONS FOR THE FOUR EXPERIMENTAL GROUPS

Problem Number	The Problem	Type of Problem	Length of Delay in sec. for:			
			10-sec. Group	20-sec. Group	30-sec. Group	60-sec. Group
1	13 29 3 get 10	Control critical problem	—	—	—	—
2	17 37 6 get 8	Example	—	—	—	—
3	31 61 4 get 22	Set Problems	10	10	10	10
4	10 39 4 get 21		10	10	10	10
5	50 102 3 get 46		10	10	10	10
6	23 49 3 get 20	Critical Problems	10	10	10	10
7	11 25 3 get 8		10	20	30	60
8	12 29 5 get 7		10	20	30	60
9	10 23 3 get 7		10	20	30	60
10	11 27 5 get 6		10	20	30	60
11	20 47 7 get 13		10	20	30	60

nated from the experiment, thereby equating all the remaining *Ss* in all four experimental groups for their ability to solve a control problem by the simple method when such a problem *precedes* the set problems.

Problem 2 is an example problem which *E* demonstrated on the blackboard.

Problems 3, 4, and 5 are set problems capable only of the complex (B − A − 2C) solution. All *Ss* in all groups who failed to solve these problems, either because of lack of comprehension of instructions or because of a lack of arithmetical ability, were eliminated from the experiment, thereby equating all the remaining *Ss* in all experimental groups for their ability to solve such problems. All experimental groups were subjected to a 10-sec. delay on each of these problems during which time they were supposed to look at the problem on the blackboard but were not allowed to write down their solutions. The *Ss*, of course, were never told how long the delay period would be. The purpose of these 10-sec. delay periods was to train the *Ss* to anticipate delays in writing down their solutions on subsequent problems, thereby enabling *E* to introduce differential delay periods for the four experimental groups on the critical problems without further interrupting the experiment.

Problems 6 through 11 were critical problems. Solutions of these problems by the set method, where a simpler method is also available, will be taken as indications of rigidity. On Problem 6, the first critical problem, the *Ss* in all groups were still subjected to a 10-sec. delay period. The purpose in giving this problem was to equate the four experimental groups by eliminating all subjects who solved this problem non-rigidly, since it would be impossible to determine further for such *Ss* the effect of different delay periods on decreases in rigidity and concrete-thinking. The remaining *Ss* in all groups were, therefore, equated in that they *all* solved Problem 6 rigidly.

The result of the equating procedure described above was to eliminate relevant individual differences both within and between groups with respect to their responses on Problems 1 through 6. The *Ss* all solved the control critical problem *non-rigidly* (Problem 1), they all solved the set problems without mistakes (Problems 3, 4, and 5) and they all solved the first critical problem *rigidly* (Problem 6). This equating of *Ss* within and between groups, however, was achieved only at the expense of eliminating 43 percent of all *Ss* tested, with approximately equal proportions being eliminated in each of the four experimental groups.

The *Ss* retained within each experimental groups were finally equated (post-experimentally) for intelligence. Scores were available for almost all the *Ss* on the American Council on Education Test. The mean A.C.E. scores were not found to differ significantly from each other.

The experimental variations were introduced with Problem 7 and continued through Problem 11. The 10-sec. Delay Group continued to solve Problems 7–11 under a 10-sec. delay condition for each problem. The 20-sec. Delay Group, the 30-sec. Delay Group and the 60-sec. Delay Group, on Problems 7–11, were subjected to delays of 20, 30, and 60 sec. per problem, respectively. It may be assumed that the greater the enforced delay, during which time the *Ss* were prevented from writing down their solutions, the greater is the time available for the perception of the problem.

Before presenting the results one further comment is necessary. The present hypothesis is that an increase in time available to *perceive* will effect decreases in rigidity and concreteness of thinking. It is therefore necessary to ensure that only the factor of time-available-to-perceive is experimentally varied and not the total length of time that each group has to solve each problem. This was accomplished by allowing two min. for each problem in all four experimental groups. Thus, on Problem 7, for example, the subjects in the 20-sec. Delay Group were made to wait 20 sec. before they could start writing their solutions and were given an additional 100 sec. to finish solving the problem while the subjects in the 30-sec. Delay Group were given a 30-sec. delay and an additional 90 sec. to finish, etc.

Results

The four experimental groups may first be compared for decreases in rigidity of solutions to the critical problems. The results are shown graphically in Fig. 1. It will be noted that on Problem 6, the first critical problem, 100 percent of all *Ss* in the four experimental groups solved this problem rigidly. It will be seen further that for all groups there are gradual decreases in the frequency of rigid solutions as they proceeded from one critical problem to the next. On Problem 7, where experimental variations in time availability are introduced for the first time, slight differences in rigidity emerge which are in accordance with the hypothesis presented. These differences, in gen-

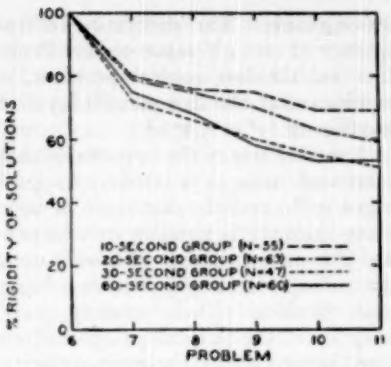


FIG. 1. Rigidity of solutions as a function of delay of response

eral, tend to become increasingly accentuated on successive critical problems. It would appear that the longer the various groups are subjected to the experimental variation in delay conditions the more different the groups become with respect to rigidity of behavior.

The 10-sec. Delay Group manifests the greatest amount of rigidity on all five critical problems and shows the smallest rate of decrease in rigidity. The 20-sec. Delay Group, on all critical problems, exhibits somewhat less rigidity and shows a somewhat faster rate of decline in rigid solutions than does the 10-sec. Delay Group. The 30- and 60-sec. Delay Groups were less rigid than either the 20-sec. or the 10-sec. Delay Groups.

The fact that recovery from the rigid pattern of problem-solving is practically identical in the 30-sec. and 60-sec. Delay Groups is worth noting. This probably indicates that rigidity as a function of time availability to perceive is not a straight-line function but levels off at about 30 sec. This means that an enforced delay of 60 sec. is no more effective in increasing perception time and in helping one to recover from rigidity than is an enforced delay of 30 sec.

Significance of the differences among the experimental groups was determined by first combining the data obtained for the 30-sec. and 60-sec. Delay Groups and then comparing the proportion of *Ss* within this combined group responding rigidly on each critical problem with the proportion of *Ss* within the 10-sec. Delay Group responding in like manner. The results are significant at the five percent level of confidence for Problem 9 and for Problem 11 ($\chi^2 = 3.84$ and 4.21, respectively).[1]

Attention may next be directed to the results bearing on the hypothesis that increases in time availability will also effect decreases in concrete thinking. A rigorous test of this hypotheses can be best obtained (1) by holding the rigidity variable constant and (2) by further equating the *Ss* within and between experimental groups for initial use of concrete tools on the first critical problem (Problem 6). To satisfy the first condition, therefore, it is necessary to compare only those *Ss* in each experimental group who solved *all critical problems rigidly*. To satisfy the second condition it is also necessary to compare only those *Ss* in each experimental group who utilized concrete aids on the first critical problem. Those *Ss* in each group who did not satisfy these con-

[1] In the opinion of the writer, the confidence levels reported are probably too conservative when the experimental results as a whole are to be evaluated for significance. This opinion is based on the following considerations: (1) They do not take into account the fact that the results of the 20-sec. Group fall intermediate between those of the 10-sec. Group and those of the 30–60-sec. Group. (2) They do not take into account the fact that although the differences between the 10-second and 30–60-sec. Groups are not significant for Problems 7, 8, and 10, nevertheless, they are in the expected direction. (3) They do not take into account the fact that the differences in rigidity among the groups become increasingly accentuated on successive problems.

ditions were eliminated from further consideration. Fig. 2 shows the results found with respect to the use of scratch paper by the various experimental groups.

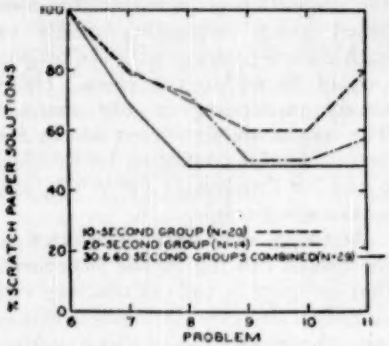


FIG. 2. The use of scratch paper as a function of delay of response

The results on the frequency of use of scratch paper, for each problem, were again found to be quite similar for the 30-sec. and 60-sec. Delay Groups and were, therefore, combined. On Problem 6, the first critical problem, 100 percent of the subjects in all groups not only solved the problem rigidly but also utilized scratch paper as a problem-solving aid. The frequency of the use of scratch paper generally decreases until Problems 8 or 9 and then levels off with the exception that there occur unexplained increases in the use of scratch paper for the 10-sec. and 20-sec. Delay Groups on the last critical problem.

The 10-sec. Delay Group on the whole used the most scratch paper. The 20-sec. Delay Group used less scratch paper than the 10-sec. Delay Group. The 30-sec. and 60-sec. Delay Groups combined used the least amount of scratch paper. These results are in accordance with the hypothesis that the greater the time availability to perceive the less the need to rely on concrete problem-solving aids. The difference in frequency of scratch paper use for Problem 11, the last critical problem, is significant at the five percent level of confidence ($\chi^2 = 5.13$).[2]

A second test of the hypothesis that increased time availability to perceive will result in decreases in concrete-thinking is possible by comparing the mean number of words used by the experimental groups in solving each problem. Unfortunately, only very few of the *Ss* in all groups verbalized their solutions, the great majority writing down their solutions in arithmetical form. Thus, it was not possible, because of the extremely small number of *Ss* in each group to start with, to equate the various groups for rigidity and for initial levels of verbalization. The meager data which were available, however, suggest that increases in time availability effect decreases in the number of words used in writing down the solutions to the problems.

DISCUSSION

The results seem to show that increases in time availability to perceive produce the hypothesized decreases in rigidity and concreteness. By creating the experimental conditions in which the *Ss* in different groups are provided with increasing opportunities to perceive it is possible to effect changes in the breadth of the 'cognitive map' (**26**), modes of thought and, consequently, end-solutions. In Tolman's (**25**) and Muenzinger's (**19**) terminology, the *Ss* in the different groups were provided with increasing opportunities to exhibit vicarious trial-and-error behavior in much the same way as the rat who hesitates at the choice point and looks alternately

[2] This statement of statistical probability is also considered to be too conservative. The reasons are similar to those given in footnote 1.

to the left and right as if to see what the alternatives are.[3]

It must not be assumed, however, that all *Ss* took advantage of this 'perceptual opportunity.' In spite of the differences which emerged among the experimental groups, over 50 percent of the subjects in all groups never recovered from their rigid pattern of behavior and an almost equal number never adopted more abstract modes of thought. Undoubtedly, uncontrolled personality factors were also operating to prevent a broad perception of the problems. Thus, it was observed that certain *Ss* repeatedly attempted to 'jump the gun' and start writing their solutions before the experimentally-specified delay period had elapsed. Others, it was observed, looked at *E* rather than at the problems on the blackboard, as if anxiously awaiting his 'Go' signal. Still others left the field completely, looking out the window, for example, rather than at the problems on the blackboard.

The introspections of the *Ss*, obtained at the end of the experiments, further revealed consistent and statistically significant relationships between 'going-out-of-the-field' behavior and rigidity. Those *Ss* who reported that they were also doing or thinking of things unrelated to the experiment were, on the whole, more rigid on all critical problems than those *Ss* who reported that they were thinking of the problem, the purpose of the experiment, etc. Examples of 'going-out-of-the-field' comments are: "Making plans for food, drink, and transportation for party Saturday night"; "Wondering if it would rain—as I had just shined my car and hoping it wouldn't"; "Tried to memorize some phonetic symbols. Counted the holes in the squares in the ceiling. Counted the venetian blind slats"; "I doodled on the blue-book, thought of home and some letters I owed and was owed"; "Thinking about many abstract things—giving very little thought to the experiment"; "I was thinking about what kind of grade I am going to get out of this course"; "I was thinking about getting a parking ticket"; "Cleaning my fingernails"; "Thinking of my future plans, this summer and next year"; "Read the State News"; "I was thinking of my fiance"; "Thinking about going to dinner and a show this evening and the events of last night."

Examples of task-oriented comments given by the *Ss* are: "I was thinking whether the next question would be of a different type"; "Trying to figure out your objective in conducting experiment"; "I was wondering what the experiment was trying to prove, how I would come out on it and bawling myself out for being misled into several foolish mistakes"; "Wondering what it was testing and why I neglected the obvious answer in one of the problems"; "Thinking of what the results would indicate"; "I was fuming silently over the fact that I had allowed myself to be taken in on the first few problems by falling victim to a problem solving set"; "I was figuring out the different possibilities of what the test was about."

In short at least two important sets of variables seemed to operate to produce the results found: (1) Personality factors which prevented broad cognitions in many *Ss* in all experimental groups and (2) structural factors, experimentally introduced,

[3] That time-to-perceive is a determiner not only of recovery from rigidity but also of learning efficiency is indicated in an experiment by Muenzinger and Fletcher (20). They found that an enforced delay at a choice point on a T-shaped discrimination box resulted in increased learning efficiency.

which resulted in the differences reported between the groups.

That the differences found among the groups are attributable to the time-available-to-perceive variable rather than to other variables is reasonable to suppose in view of the fact that (1) all *Ss* within and between groups were equated such that no individual differences existed in their solutions of Problems 1 through 6, (2) all *Ss* in all groups had a total of two min. per problem and (3) the groups were not significantly different from each other in mean intelligence scores (as measured by the American Council on Education Test).

The fact that the results for the 30-sec. and the 60-sec. Delay Groups were almost identical for both rigidity in problem-solving and for concreteness of thinking may be attributed to the levelling-off effect taking place at about 30 sec. Delay time beyond this is apparently spent in non-problem-solving activity rather than in pre-solution perceputal activity. This does not imply, however, that delays up to 30 sec. are necessarily spent in perceptual activity. Written comments obtained from the *Ss* after the experiment indicated that numerous subjects became 'perceptually satiated' before 30-sec. and, for that matter, before 20 sec. and even before 10 sec. In some cases the comments revealed this was due to the fact that they perceived both alternative solutions and then left the field. In other cases 'perceptual satiation' occurred relatively soon, thereby preventing the perception of the more simple solution.

The theoretical considerations on which the present experiment is based and the results found lead to certain implications for further research concerning differences between rigid and flexible personalities. Other things being equal, it may be hypothesized, differences between persons characterized as rigid and others characterized as less rigid may be attributable, to some extent at least, to personality differences in time availability. It is being suggested here that while in the present study time availability was treated primarily as a structural determinant of differences between experimental groups it may also be regarded as a primarily functional determinant of differences between rigid and flexible personalities. Time availability, if present to a large degree within the personality structure, makes possible broader cognitions, more abstract thinking (more vicarious-trial-and-error), and, consequently, greater flexibility of behavior. Conversely, if time availability as a personality variable is present to a lesser degree cognitions are generally narrower, thinking is more concrete, and behavior, consequently, is more rigid.

Similarly, to the extent that rigidity is related to prejudice, as is indicated by the research of Frenkel-Brunswik (7, 8) and Rokeach (22), it may also be hypothesized that personality differences in time availability exist between prejudiced and unprejudiced personalities.

The results found in the present study suggest several further problems for future research. It would appear that the concepts of impulsiveness (1, 5), impulse delay (21), and delay reactivity (6) are related to the present concepts of time availability and rigidity. Further research is necessary to determine to what extent this is so. Another problem of interest is whether compulsive-obsessiveness, "which appears from without as ability to wait" (1, p. 12), is related to the concepts used here.

While it is not possible at the present time to go into any great detail about possible dynamic determiners of time availability it may be suggested here that some individuals, because of past experiences with frustrating situations involving delay of need satisfaction, become generally incapable of tolerating frustrating situations. To allay anxiety, such individuals learn to react relatively quickly to new problems which continue

to arise. This results in consistently narrow perceptual organizations. At the same time it also prevents the establishment of abstract modes of thought since abstract thinking can occur only when there is time available to think through a problem. The inevitable consequence, it is hypothesized, is behavioral rigidity.

Summary and Conclusions

The present experiment arose from an earlier study of the relationship among prejudice, rigidity, and concrete-thinking (22). The hypotheses tested are that increases in time availability to *perceive* will result in (1) decreases in rigidity and (2) in decreases in concrete thinking.

Rigidity was measured by an arithmetical technique in which critical problems may be solved rigidly or non-rigidly. Concreteness of thinking was measured primarily by the frequency of use of scratch paper as a problem-solving aid.

Four experimental groups, subjected to delays of 10, 20, 30, and 60 sec., respectively, had increasing opportunities to perceive the problems presented. These groups were equated with each other with respect to intelligence, initial level of rigidity, time exposed to each problem, ability to solve the problems presented both by the rigid and non-rigid methods, and for comprehension of instructions. In comparing the experimental groups for concrete-thinking the *Ss* within and between groups were further equated for rigidity on all problems and for initial level of concreteness.

The main results were:

1. The 10-sec. Delay Group, having the least opportunity to perceive the problems, gave the greatest frequency of rigid solutions and utilized concrete aids more frequently than the other groups.

2. The 20-sec. Delay Group gave somewhat fewer rigid solutions and utilized concrete aids less frequently than did the 10-sec. Delay Group.

3. The 30-sec. and 60-sec. Delay Groups, having the greatest opportunity to perceive the problems, manifested the smallest frequency of rigid solutions and utilized concrete aids least often.

4. The differences among the experimental groups tended to become increasingly greater with respect to rigidity as they progressed from one critical problem to the next.

5. No differences appeared between the 30-sec. and 60-sec. Delay Groups with respect to the rigidity and the concrete-thinking measures used. These results indicate that rigidity and concreteness as a function of time availability, at least with the particular problems used, levels off at about 30 sec.

6. Certain theoretical implications of the results are discussed which bear on the problem of personality differences between rigid and non-rigid individuals and, to a lesser extent, between high-prejudiced and low-prejudiced individuals.

(Manuscript received May 19, 1949)

References

1. Bergler, E. On the psychoanalysis of the ability to wait and of impatience. *Psychoanal. Rev.*, 1939, **26**, 11–32.
2. Buytendijk, F. J. An experimental investigation of the influence of cortical lesions on the behavior of rats. *Arch. neerl. Physiol.*, 1932, **17**, 370–435.
3. Cameron, N. Cerebral destruction in its relation to maze learning. *Psychol. Monogr.*, 1928, **39**, No. 1.
4. Cotton, C. B. A study of the reactions of spastic children to certain test situations. *J. genet. Psychol.*, 1941, **58**, 27–40.
5. Fenichel, O. *The psychoanalytic theory of neurosis.* New York: Norton, 1945.
6. Foster, A. Factorial studies in personality. (To be published)

7. Frenkel-Brunswik, E. Dynamic and cognitive categorization of qualitative material: II. Application to interviews with the ethnically prejudiced. *J. Psychol.*, 1948, 25, 261–277.
8. Frenkel-Brunswik, E. Tolerance toward ambiguity as a personality variable. *Amer. Psychologist*, 1948, 3, 268. (Abstract)
9. Goldstein, K. Concerning rigidity. *Character & Pers.*, 1943, 11, 209–226.
10. Goldstein, K., & Sheerer, M. Abstract and concrete behavior. An experimental study with special tests. *Psychol. Monogr.*, 1941, 33, No. 2.
11. Hamilton, J. A., & Ellis, W. D. Behavior constancy in rats. *J. genet. Psychol.*, 1933, 42, 120–139.
12. Kounin, J. S. Experimental studies of rigidity. *Character & Pers.*, 1941, 9, 251–272; 273–282.
13. Krechevsky, I. Brain mechanisms and variability. *J. comp. Psychol.*, 1937, 23, 121–138; 139–163; 351–364.
14. Krechevsky, I., & Honzik, C. H. Fixation in the rat. *Univ. Calif. Publ. Psychol.*, 1932, 6, 13–26.
15. Lewin, K. *A dynamic theory of personality.* New York: McGraw-Hill, 1935.
16. Luchins, A. S. Mechanization in problem solving: the effect of Einstellung. *Psychol. Monogr.*, 1942, 54, No. 6.
17. Maier, N. R. F. The effect of cerebral destruction on reasoning and learning. *J. comp. Neurol.*, 1932, 54, 45–75.
18. McAndrew, H. Rigidity and isolation: A study of the deaf and blind. *J. abnorm. soc. Psychol.*, 1948, 43, 476–494.
19. Muenzinger, K. F. Vicarious trial-and-error at a point of choice. A general survey of its relation to learning efficiency. *J. genet. Psychol.*, 1938, 53, 75–86.
20. Muenzinger, K. F., & Fletcher, F. M. Motivation in learning. VII. The effect of an enforced delay at the point of choice in the visual discrimination habit. *J. comp. Psychol.*, 1937, 23, 383–92.
21. Rappaport, D. *Diagnostic psychological testing.* Vol. II. Chicago: Year Book, 1946.
22. Rokeach, M. Generalized mental rigidity as a factor in ethnocentrism. *J. abnorm. soc. Psychol.*, 1948, 43, 259–278.
23. Strauss, A. A., & Werner, H. Experimental analysis of the clinical symptom 'perseveration' in mentally retarded children. *Amer. J. ment. Def.*, 1942, 47, 185–188.
24. Strauss, A. A., & Werner, H. Comparative psychopathology of the brain-injured child and the traumatic brain-injured adult. *Amer. J. Psychiat.*, 1943, 99, 835–838.
25. Tolman, E. C. The determiners of behavior at a choice point. *Psychol. Rev.*, 1938, 45, 1–41.
26. Tolman, E. C. Cognitive maps in rats and men. *Psychol. Rev.*, 1948, 55, 189–208.
27. Vigotsky, L. S. Thought in schizophrenia. (Trans. by J. Kasanin). *Arch. Neurol. Psychiat., Chicago*, 1934, 31, 1063–1077.
28. Wekner, H. Abnormal and subnormal rigidity. *J. abnorm. soc. Psychol.*, 1946, 41, 15–24.
29. Werner, H. The concept of rigidity: A critical evaluation. *Psychol. Rev.*, 1946, 53, 43–52.
30. Werner, H., & Strauss, A. A. Causal factors in low performance. *Amer. J. ment. Def.*, 1940, 45, 213–218.

ANALYSIS OF DETOUR BEHAVIOR: IV. CONGRUENT AND INCONGRUENT DETOUR BEHAVIOR IN CATS

BY PAUL H. SCHILLER †

Yerkes Laboratories of Primate Biology, Inc.

Statement of the Problem

If an organism is offered an opportunity for need-reduction whenever it performs a certain locomotor act, this act tends to be displayed repeatedly, under similar circumstances, whenever the need exceeds a certain level. The act acquires the function of being conducive to the chance of reward; it becomes a routine. From the viewpoint of the experimenter, operationally, this procedure is called instrumental conditioning, and from the viewpoint of the animal, anthropomorphically, problem-solving behavior.

Two essentially different types of instrumental conditioning are sometimes discriminated. In the one case an act can be required that is relevant to the situation to be produced by the act, while in the other case the act is irrelevant to the situation. Thorndike (17) called these the congruent and the incongruent acts. He found that his cats would learn to lick or scratch themselves 'in order to' get out of a box as readily as to manipulate a lever, button, or string. Now, the autistic gesture required has nothing to do with the releasing mechanism, whereas the lever, etc., actually release the door. Therefore, the object-manipulation is congruent with its consequences, while the self-manipulation is an incongruent act. There is some indication in Thorndike's illustrative data that incongruent acts take more time to become established as routines than do simple object manipulations, and require less time than the complex operations of combined latches.

To be sure, both types of situation involve acts that are, for the cat, consequential. The difference is that in the congruous or relevant case the 'reward' or 'goal' is an 'inherent' consequence, in the incongruous or irrelevant case it is an 'arbitrary' one. The inherent consequence is produced by object-relationships, while the arbitrary one is based on human interference.

Critics of Thorndike's studies felt that both situations required, from the cat, equally irrelevant or incongruent acts. After all, the cat could see the relationship between lever and door no more than it could see the relationship between self-licking and release. The hidden mechanism must appear as magical to the cat as the whim of the experimenter. Gestalt-psychologists suggested the study of truly congruent or relevant situations and giving the animal a fair chance to perceive relationships and to behave insightfully. Adams (1), who repeated the Thorndike experiments with the greatest care, found that better adjusted cats behave more sensibly. Yet, even in his own problems, with the relevant relationships made visible, accidental success was necessary. The final act, conducive to success, be it rotating of a lever, pulling a string or a box, tended to be repeated, much in the same way as the stereotypes studied by Guthrie and Horton (6). The Guthrie cats

† Editor's note: Dr. Schiller was killed in a skiing accident soon after he had submitted this paper for publication. Dr. Henry W. Nissen has kindly read the galley proofs.

made few errors because an accidental range of activities was equally conducive to release and food; all they had to do was to touch the pole. It is hard to make a congruous task that simple. The clearest of Adams' situations was the detour problem that required the following of a roundabout pathway in familiar surroundings.

Detour problems can produce *par excellence* relevant or congruent solutions: the circumvention of an obstacle is an adjustment to the geometrical qualities of the situation by a locomotor pattern, leading to a certain locus located by perception of a bait and the obstacles between animal and bait. Adams' cats did very well in varied detour situations. The advantage of such a situation over the manipulation of hidden mechanisms seems to be obvious. Since differences in the manipulanda, however irrelevant to the human observer, influence the rate of learning considerably, a comparison of congruent and incongruent acts is extremely dubious. Such a test must certainly utilize strictly comparable effector activities.

Thorndike's result might be based on the fact that a simple conspicuous visual object releases manipulative response more readily, previous to specific learning, than confinement in a box induces autistic gestures. The differential availability of the responses might blur the factor of congruency or relevancy of its appearance (later called belongingness by Thorndike). To test the variable of belongingness the availability must be kept constant.

The plan of the present experiment was developed out of this dilemma. Adaptation to simple detours should be compared with a similar locomotor pattern guided, however, by irrelevant rather than relevant perceptual cues. Such a comparison still will not be entirely fair: in the general day by day experience of the cat, things of interest appear to be contacted usually in a place where they were sensed by distance receptors and not systematically in a different place. Yet, if we disregard the probable generalizations of the 'consistency of reality,' in the experimental situation proper we can provide the animal with strictly comparable tasks of a relevant and irrelevant nature, respectively. Let a cat, before or after it has learned to use a U-shaped detour, in various positions, *toward* the food, learn to make another detour, not toward the food seen and smelled, but always in the *opposite* direction. The locomotion will be essentially the same: three turns of about 90° in succession. To be sure, both detours, the relevant as well as the irrelevant one, are given in randomly alternating spatial directions. Place-learning is excluded and response-learning required (cf. Tolman *et al.*, 18).

Procedure

A detour apparatus developed by the author (12, 13) for the study of delayed detour performances in aquatic animals was adapted for giving alternate detour-problems to cats. An alley is built of boxes. The middle branch is the starting chamber with wide-mesh wire windows on both sides and a slide-door in the middle, opposite to the entrance into the starting chamber. The slide-door opens to the choice compartment, a cross-section which connects both side branches. The ends of the side branches are the feeding boxes behind the wire windows of the starting chamber. If food is placed behind one of the windows, the cat perceives which of the two boxes is baited and, passing through the opened doorway, can make a choice to turn either to the left or right, describing a U-formed roundabout path, in either case. The detour is three to four times the body length of the cat and requires about three or four sec. to traverse it. The delay involved is usually not more than one or two sec. Since Adams has shown successful delayed response for hours, there was no point in making the alleys transparent. It might be of interest, however, to repeat the experiment under

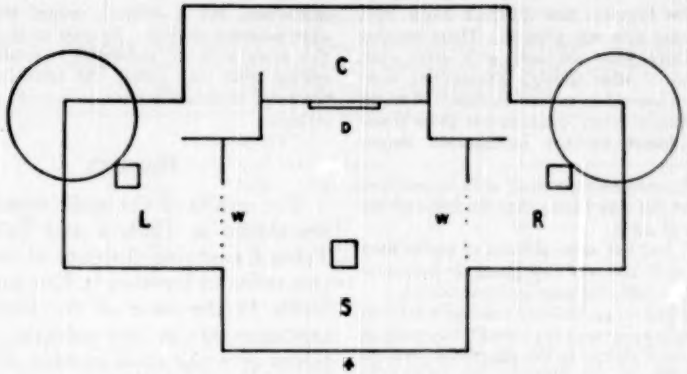


FIG. 1. Ground plan of alternative detour apparatus for cats

S = starting box, C = choice box, L, R = left and right feeding chambers, w = windows, d = slide-door. The squares show the projections of the openings for dropping food, the circles the rotating discs to reverse position of food. The animal is placed in and observed from the point marked by the arrow.

such conditions. In fish it was found (15) that an obstacle that partly hid the food to vision was more easily adapted to than an all-transparent partition wall. Fig. 1 shows the ground-plan of the boxes.

In this apparatus two types of training were given. The act relevant to or congruent with the situation was a detour, to the left or right, *toward* the baited place; the act irrelevant to or incongruent with the placement of the perceived lure was a detour in the *opposite* direction, away from the baited feeding chamber, to the unbaited one. Each type of problem was trained until the criterion of 20 consecutive errorless trials was reached. Thereafter, 20 critical trials were given; they consisted of situations in which both chambers were actually baited (but bait was shown only on one side).

In the relevant detour situation the food was placed and left in the feeding chamber all the time. In the irrelevant problem situation the food, after being presented, was withdrawn at the moment the animal approached the slide-door and simultaneously with the opening of the latter food was placed in the opposite feeding box. This was performed by the turning of two discs, one rotating the food in an arc of 90° into the box, the other one rotating it out of the other box. This operation produced a noise, which for human ears was equal on both sides. Eventual inequalities were neutralized by the fact that both sides were used equally often. In critical trials this noise was omitted, the disc with the food left inside, and the correct side loaded after the accomplished choice.

To get the animal back to the starting place, the younger animals had the side window opened releasing them straight back to the starting place. The older animals were required to retrace the detour. This latter technique was employed in the last two series with the younger animals also, and resulted in no noticeable change in performance. In both cases the exit was rewarded and the next trial started as soon as the food at the starting place was consumed. A trial started with placement of food behind the window, so that it provided acoustic, optic, and olfactory stimuli at the same time. If there was a delay in administration of food, the animals sniffed or pawed at one of the windows, or both. Trials were repeated in series up to 20 or 30, according to the willingness of the animal to work. A session lasted usually 12 to 16 min. The number of responses was recorded per time unit of two min.

Both types of problem were given to every animal except one, but the order of reversals was varied. Five cats were used. Three of them, D, B, A, all males, were raised at home since weaning age, and were subjected to these experiments at the age of 8 to 16, 10 to 20 weeks, and 6 to 9 months, respectively. E and C, a male and a female from the neighborhood, were used at the approximate age of 4 to 5 and 10 to 12 months.

Of these cats, A alone had extensive earlier training in a stereotyped pattern of entering one of two boxes and returning, reinforced by feeding after he returned. Entering both of the left and of the right side was developed. Thereafter,

alternation between two different boxes both to the same side was trained. These learning performances were obtained with great ease. No double or other multiple alternations, however, were learned in several hundreds of trials. The congruous detour problem was given thereafter, followed by the incongruous detour problem.

Cat B made only 60 trials with stereotyped turns into the side boxes; then he followed the sequence of cat A.

Cat C had the same amount of preliminary training as B, and was then given the incongruous problem first, the congruous one second.

Cat D had no preliminary training whatever, started right away with the incongruous problem and was then shifted to the congruous problem situation, like C.

Cat E had no preliminary training either. He was given the congruous detour problem first and then shifted to pseudo-detour problems with the food *not* presented before choice had been completed.

Both cats B and D had their problems reversed twice after accomplishing the above mentioned tasks. Both of them thus had worked in four series, alternating the congruous and the incongruous situations in reversed order of sequence. Furthermore, they were tested after completion of the main experiment with the same patterns of movement without any food incentive presented, but released as soon as they oriented to a window and rewarded, after choice, either on the same or the opposite side from where they started. Thus, pseudo-congruous and pseudo-incongruous problems were given, in order to see whether stereotypes of postural patterns are sufficient to maintain or relearn the earlier habits.

The older animals were fed only in the experiment, which took place in the author's home at night. The younger animals received some additional food in the morning, 12 hours before the experiment. After the experimental session each cat received milk. In the experiment Calo Cat Food was given in pieces of $\frac{1}{4}$ to $\frac{1}{2}$ oz. per trial. Experiments were run daily with the exception of a few Saturdays or Sundays.

In the initial training period both sides were given in blocks of trials first to the one, then to the other side, in such a way as to break side preferences. After this, regular alternation was introduced and kept on until 10 consecutive trials occurred with no error. This was the case after 40 to 90 trials in any one series. After this, a standard pattern of systematic random alternation was adhered to. In all training trials the correction method was adopted: the cat was allowed to return to the choice point and go over to the actually baited feeding box. Thus, no immediate, but a delayed, reward was given after incorrect choices. As soon as 20 consecutive trials were all successful (see tables), 20 critical trials were given. All critical trials of the main experiment were 90 percent or more correct.

Results

The results of the main experiment are shown in Table I and Table II. Table I contains the data of the first two series of learning in four animals, Table II the data of the additional two reversals in two animals. Both tables give the total number of trials necessary to reach the criterion of 20 consecutive successful trials. These latter trials are included but not the critical trials. The number of errors is given too, from which the number of immediate reinforcements can easily be computed by subtracting the errors from the trials. The time taken to complete the training in several sessions is totaled for every series. The number of days can be obtained, by dividing the time by about 12 or 16, for the younger and older animals, respectively.

It is to be seen that the incongruent act was, in terms of trials, more difficult to learn for A and B who learned it as their second task; it was, however, easier for C and D who made it in their first series. The differences, of course, are void of statistical significance, since the individual variation in the rate of learning is quite considerable. Yet the direction of differences is consistent and suggests that the second, reversed problem, whichever it was, took more trials to be mastered than the first one.

The same is true, moreover, if the number of errors is considered, though the total number of errors is greater in the incongruous problem (and, consequently, correct responses are fewer).

Regarding the time scores, in three

TABLE I

TIME IN MINUTES, AND NUMBER OF TRIALS AND ERRORS IN LEARNING CONGRUENT AND INCONGRUENT ACTS THAT REQUIRE COMPARABLE EFFECTOR ACTIVITIES

Roman numerals indicate order of sequence of tasks.

Cat		Congruent				Incongruent		
		Time	Trials	Errors		Time	Trials	Errors
A	I	108	223	47	II	220	259	117
B	I	120	180	57	II	180	218	81
C	II	202	215	89	I	196	170	62
D	II	104	168	55	I	144	145	43

animals the incongruous learning required more time not only in total, but per trial, too, and this latter is true of the fourth animal, C, as well. This fact does not mean, however, that more reinforcement is necessary to master the incongruous problem. As a matter of fact, the total number of the immediate, non-correction reinforcements is somewhat less in the incongruous case. These trials are spaced further apart in time than the congruous ones. This lower rate of responding suggests that there is an increased hesitance in the case of the incongruent as compared to the congruent problem-solving behavior. Such an *emotional* difficulty, nevertheless, does not interfere with *cognitive* success.

This analysis shows that *there is no dramatic advantage* of the congruous task over the incongruous task in situations that require comparable effector activities.

The two younger cats showed in their further series of reversed learning no indication of any preference in relearning either task. The second reversal produced in the third task (congruous for B, incongruous for D) definite improvement which was kept up, after third reversal, in the fourth series.[1] It corroborates the finding that neither of the two tasks has any clear-cut advantage as an easier one over the other task. Total time required and errors are still greater in the incongruous task than in the congruous one. Table II shows the data together with a sum-total of all data. It is a consistent feature in all series that the differences in trials, errors, (and reinforcements) are so small that the samples of the two groups could have been taken from the same population.

The time difference alone seems to be of some significance. It is at the level of $P = .1$ (including non-independent data of Table II). Since the time recorded included feeding as well as returning time, it expresses the general rate of work. Cumulative learning curves plotted to the time show that it is the extinction of the congruent act that is initially slower than the extinction of the incongruent act. Most of the hesitation accompanies errors.[2]

The congruous task was performed with a mean speed of 39 sec. per trial, the incongruous one with 53 sec. per trial ($PE = 3.2$). Such a difference in the rate of work is suggestive of emotional differences in the two situations. Moreover, it might be a hint that, after all, the two locomotor pat-

[1] The increased speed of learning after repeated reversals, but not the first, has been called the learning of re-learning (Buytendijk, 3).

[2] A similar finding in the octopus is reported elsewhere (13).

TABLE II

TIME IN MINUTES, AND NUMBER OF TRIALS AND ERRORS IN RE-LEARNING CONGRUENT AND INCONGRUENT ACTS. Roman numerals indicate order of sequence. Data of both tables are added in last row.

Cat		Congruent				Incongruent		
		Time	Trials	Errors		Time	Trials	Errors
B	III	60	112	21	IV	72	116	35
D	IV	58	103	28	III	64	101	42
Total		652	1001	297 (704)*		876	1009	380 (629)*

* Number of immediate reinforcements is given in parentheses. The *t*-values of significance are 1.83 for time, .14 for trials, 1.27 for errors.

terns were not entirely equal in their ease of performance. Maybe the postural sequence was learned as such, rather than as a response to food. Considering the fact that the animal, at return, is facing the entrance of the starting box and has its rear at the slide door, the turn toward the window has to be added to the total response pattern. In a congruent act, then, the turn to the window is continued to the door where it is followed by an opposite double turn of U-shape, so that the actual pathway is S-shaped. In the congruent act, the animal continues in the doorway the same sense of turn all around, so that its whole path is actually circular. Which one has greater probability must be tested.

It thus appeared necessary to check upon the animal's attention or orientational responses to the lure. An actual test was therefore carried out that required either type of postural sequences with *no* incentive presented at the start.

Following the critical trials of the fourth series, the two kittens were subjected to a further reversal, by opening the door as soon as they oriented to an empty window and reinforcing the congruent act in B, the incongruent act in D. B did not learn the pseudo-congruent act at all. In 200 trials he reached only 35 percent correct. D, on the other hand, learned the pseudo-incongruent act readily (in which B persisted). A further reversal showed that B now learned the pseudo-incongruent act readily, but D failed to learn the pseudo-congruent act. D reached only 60 percent correct in the last 40 of 200 trials. Both animals worked slower in these series than in the preceding ones, and equally in both the congruous and incongruous pseudo-task. The data are given in Table III.

To test an eventual loss of orientation in consequence of the many reversals in training after these last series, both kittens were given 20 real congruent trials with food. The earlier speed of work was at once restored. D performed 100 percent correct and B about chance (his last habit was 'incongruent'), but after 50 more trials B reached the 90 percent level.

If there was no incentive, both animals were inclined to display the circular rather than the S-shaped pathway, irrespective of the preceding training and of the actual administration of reinforcements. This peculiar preference was obtained even though the reaction potential built up in the

entire course of reversals was not made higher for the incongruent act: it will be recalled that there were a few less immediate reinforcements for the incongruent act than for the congruent one (see last row of Table II). The great availability of the incongruent act might thus be due to some unlearned preference that produced the dominance of the circular pattern.

To test this possibility, a hitherto untrained cat, E, was trained first in the real congruous situation, then without food. This kitten learned the task more readily than any other cat: it made the 20 errorless trials in a total of 77 trials (and 54 min.), with very few errors. When there was no bait, he was reinforced for the same act, turning toward the side from which he started. He now worked much slower (one trial in every 55 instead of 42 sec.), and made more mistakes, but reached the criterion of learning very soon. Shifting over to the pseudo-incongruous task took him about twice the amount of learning at about the same low speed of work. (See Table III.) The pattern of the incongruent act thus does not seem to be innately dominant.

This kitten E serves, in another respect, as a control for D, who had not had any previous training either, but learned the incongruous task first. For this he needed more than double the amount of training than E for his first, congruous, task. If we neglect the difference of tasks, it is obvious that the other cats required even more trials in their first problems. The amount of work to reach the first learning criterion is roughly proportional to the amount of preliminary training with stereotypes (A had most, B and C some, D and E none). This points to an interference of earlier habits with the present situation.[3]

It is to be seen in this context that a cat that had no experience with the incongruent act does not prefer the pseudo-incongruent pattern, whereas the cats that had such an experience prefer this pattern to the pseudo-congruent one. Why the postural habit of an incongruent act *prevails* over the other sequence of turns after the original repertoire of movement patterns has been enriched with it, is a puzzling question. S-paths and circular paths seem to have originally

[3] A verification of this possibility is given by Carmichael (4). Although his study was done with rats, the same apparatus has been used with cats also.

TABLE III

TIME IN MINUTES, AND NUMBER OF TRIALS AND ERRORS IN CHANGING OVER TO PSEUDO-CONGRUENT AND PSEUDO-INCONGRUENT ACTS WITH NO INCENTIVE PRESENTED. Roman numerals indicate order of sequence. Criterion of learning is given in the row below data of each animal.

Cat		Pseudo-Congruent				Pseudo-Incongruent		
		Time	Trials	Errors		Time	Trials	Errors
B	V	150	200 35%	143	VI	52	78 100%	14
D	VI	122	200 60%	126	V	54	89 100%	16
E	II	64	69 100%	13	III	128	145 95%	59

equal chances in young cats, as follows from general observation also.

It might be suggested that the cats were conditioned to the absence of food rather than to the presence of it. Yet this stimulus, if it existed, had not been associated more often with the incongruent than with the congruent act. In any case, unequal reaction potentials have been produced, with the only detectable difference in the timing of consecutive reinforcements.

This enigmatic behavior could be stated, of course, in cognitive terms: the animal is inclined to choose the other side if there is no food perceived on this side. In other words, it is harder to go toward an obviously non-baited side than to go away from a baited side to the other one, where food has also often been given. A set to expect food in one of the boxes was definitely developed.

The non-incentive trials were conclusive in only one respect. The cats have certainly attended to the bait in the main experiment. The congruent act was performed in a consistent and rapid way only when bait was presented. This finding enforces the claim that the main experiment was dealing with genuine detour and counter-detour situations.

Discussion

To test strict reinforcement learning versus cognitive or perceptual learning it is an important step to demonstrate that the animal readily develops stereotyped routines in consequence of coincidence of an incidental movement with a situation that is conducive to reinforcement. Using comparable effector activities, it was shown that congruent and incongruent acts are learned, in terms of trials and errors, equally readily, although the former has a stronger initial persistence while changing over to the latter than has the latter while changing over to the former.

This relative difficulty of learning the incongruent act may be interpreted as an emotional impediment due to the discrepancy of that sort of situation with the general experience of the animal. There always remains an advantage of the congruent over the incongruent situation: the learning sets developed in the daily activity of the animal favor certain activities under certain circumstances more than others. This favoring is usually in a sense that conforms to the external world as well as to the structure of the organism. This *consistency of reality* reinforces certain kinds of activities to an extent that the organism has difficulty in adapting to contradictory conditions. It is poor competition only if in the restricted realm of an experiment a set of irrelevant attitudes is artificially developed.

As a test of experimental set-effects, I am at present performing some studies on the delayed response in chimpanzees. The task is to find the bait under that one of two containers that has *not* been baited, when the other was loaded with food while the animal was watching. Two sophisticated adult animals have learned the trick in 200 trials each, one to the criterion of 95 percent, the other to 100 percent in 40 consecutive trials. This was obtained in spite of several thousands of earlier contrary experiences.

Other cases of incongruent behavior in apes have been occasionally observed. A pioneer expert, Boutan (2), compared the behavior of a gibbon with five puzzle boxes, four with visible, the fifth with a hidden latch mechanism. He reported some difficulty with the latter box but added that this box made a noise that

frightened the gibbon. Guillaume and Meyerson (5) describe a few stubborn clumsinesses of their chimpanzee artist that remind me of token responses. Some of Köhler's 'good errors' seem to fall in the same class as the 'superstitions' of Skinner (16). I have tested chimpanzees of various age groups and found that there are peculiar patterns of handling objects displayed under favorably undisturbed conditions with no reference to whatever 'goal.'

These apparently native stereotypes constitute elements of 'tool-using' (cf. 14). If equally available motor patterns are used, there appears to be no essential difference in the rate of learning and of extinction (a) to use a stick to pull in food or (b) to push that food far away with it. This equalization of a congruent or relevant act with an incongruent or irrelevant act is achieved by having the animal get a reinforcement subsequent to either act.

Since both the apes and the cats have used varied effector activities in equivalent cases of responses to left and right, it is not the response but the place that is specific in the case of correct behavior. Yet the performances studied here are not those of remembering places or directions in the sense of Tolman's and his associates' (18) recent studies. The alternatives of my animals are equivalent in their relation to external cues and their general pattern; they depend on the actual locus of orientational responses provoked by external cues first on one, then on the other side. This locomotor pattern is a response made conditional to distant or postural cues. In a lower organism, the octopus (15), thigmotactic cues seem to suffice for a delayed choice, while in the cat (and the ape) a more generalized pattern guides locomotion. This pattern is equally available to the left or the right, yet it need not be due to any perceptual schema, but can be a case of motor transfer.

Unfortunately, little work has been done in the field of response generalization, to which Lashley (9) has drawn attention. A familiar melody can be played on the piano keyboard with a fingering never practiced before. It is likely that well learned sequences of manipulative or postural responses tend to establish a locomotor pattern with many degrees of freedom allowing for minor or even major adjustment, transfer, equivalences in the effector functions. Beside the generalized discrimination sets studied by Harlow (7), the expanding sets of motor equivalents deserve to be explored. *Motor generalization patterns* may be expected to account for immediately adaptive responses, for congruent activities of insightful appearance.

To test the mechanisms of learning, the repertoire of pre-training movement patterns should be studied explicitly. The cat, e.g., has a predilection for pulling strings and pushing movable objects of certain sizes and textures, as seen in the play activities of kittens reared in isolation from fellow members of the species. It seems to me that many of Adams' impressive observations on primary problem-solution are reports on such innate motor patterns in cats. It is possible that a coincidence of native patterns and actual directions of activities in a problem situation enhances the chances of learning. Thus, kittens learn in one or two trials to pull a string, operate a lever, etc. In the present experiment, however, neither of the acts compared was found to be natively dominant over the other one. Both of them utilized the response of brushing along edges and corners.

In the rat, Lorge (10) has found no reliable difference in the ease of learning to perform activities like face-washing, scratching, begging, stand-up, and pressing four different clips.[4] But a complex latch box is certainly learned with more difficulty than a simpler one. It is suggestive to compare instrumentational problems with the congruous and incongruous detour tasks.

In the opening of latch boxes Kohts (8) has reported two radically different

[4] He unfortunately misnames the problem as one of irrelevant rewards or wants, though his animals were equally anxious at all times to get out of the box.

methods of work displayed by a monkey. Simple hasps, padlocks, etc. have been opened with great adaptation to mechanism, position, and size. There was an immediate generalization of different structures. But if two or more locks were combined, all of them known individually to the monkey, the attempts made were blind trials, confused fumbling with the locks. In this case, of course, the two situations are not comparable and hence we cannot infer that 'ideational' learning is easier than learning by mere repetition. Although the opening patterns were made equally available, they certainly were not equally promptly reinforced. The apparent difficulty of temporal organization is mainly due to the simple fact that the reinforcing state of affairs is reached with considerably more delay and variation of means in a combined than in a simple latch-problem.

It is a prerequisite to studies on learning differentials in physiological, pathological, and neurological analysis of adaptive behavior to evaluate the grade of difficulty of the problems. In order to devise the tasks as comparable, the native and the acquired availability of motor patterns should, so far as possible, be equated.

Lashley (9) has found in rats that problems easy to learn for normal animals were more difficult for animals with cortical lesions, while the normally more difficult tasks were manipulated in no worse manner by the operated rats. The interpretation is that the normals learned these complex problems by random activity, i.e., the same way as the operated animals learned all problems. The easier tasks, on the other hand, were mastered by the normal animals with the aid of attention or insight, impaired in the brain-injured rats.

Routines of motor patterns can be established by short span reinforcements, as in Skinner's or Guthrie's animals. Roberts (11), on the other hand, never found such stereotypes when the 'irrelevant' touching of a pendulum was required (from rats). The reason for his outcome is simple: he delayed the opening of the door after the required act was performed. If reward comes later, the first necessary movement has less chance to be associated and routinized, and more variation and generalization can take place. It is the last movement before termination of the situation that tends more strongly to be repeated than the others. Roberts observed that his rats never missed running to the door before touching the pendulum. The conjecture, therefore, that the rat was insightful when, after training, it used the door immediately if found open and skipped to touch the pendulum, is unwarranted. Adaptiveness is a function of the size of acquired motor units. Comprehensive units, however, are not necessarily rational.

Animals thus seem, whether primates, cats, or rats, to learn an incongruous task as readily as a congruous problem, when only the effector activities are equally available. It therefore appears to be a general feature, of mammalian learning at least, that to learn irrelevant things, though emotionally loaded, is not more difficult, cognitively, than to learn relevant connections.

Summary

1. Four male cats and one female were tested in both a congruous and an incongruous detour problem. Two groups of animals followed an opposite sequence of reversals. Both problems required a comparable pattern of simple turns.

2. The animals learned with a greater general working rate to solve the congruous problem of running around partition walls toward the food than to run around toward the opposite, apparently not baited, feeding box. The total numbers of trials necessary to reach the criterion of learning did not, however, differ reliably in the two situations.

3. The results of repeated reversals of the problems indicated that the initial differences, due to sequential factors, tended to disappear.

4. The difference in working rate is interpreted as a consequence of emotional strain rather than of a learning differential. The frustration of the incongruous situation, at variance with the general experience of the animal, delays his learning.

5. The circular pathway required in the incongruent act was prepotent in animals who had equal training of both patterns, as shown in series with no incentive presented. A naive animal, however, performed the pseudo-congruent S-shaped postural sequence with greater ease than the pseudo-incongruent act.

6. These results are compared with instrumentational and latch-box problems in mammals. Observations that experienced chimpanzees easily learn to choose the unbaited rather than the baited container in the conventional delayed response situation and learn to push away food as readily as to pull it in with a stick are suggested as parallel to the facts reported here.

(Manuscript received April 25, 1949)

References

1. Adams, D. K. Experimental study of adaptive behavior in cats. *Comp. Psychol. Monogr.*, 1929, 6, No. 27, pp. 166.
2. Boutan, L. Les deux methodes de l'enfant. *Act. Soc. Linn.*, Bordeaux, 1914, 67, 3–146.
3. Buytendijk, F. J. J. Über das Umlernen. *Arch. Néerl. Physiol.*, 1930, 15, 283–310.
4. Carmichael, L. Learning which modifies an animal's subsequent capacity for learning. *J. genet. Psychol.*, 1938, 52, 159–163.
5. Guillaume, P., & Meyerson, I. Recherches sur l'usage de l'instrument chez les singes, I–III. *J. de Psychol.*, 1930, 27, 177–236; 1931, 28, 481–555; 1934, 31, 497–554.
6. Guthrie, E. R., & Horton, G. P. *Cats in a puzzle box.* New York: Rinehart, 1946.
7. Harlow, H. F. The formation of learning sets. *Psychol. Rev.*, 1949, 56, 51–65.
8. Kohts, N. Les aptitudes motrices adaptives du singe inferieur. *J. de Psychol.*, 1930, 27, 412–447.
9. Lashley, K. S. Studies of cerebral function in learning: XI. The behavior of the rat in latch-box situations. *Comp. Psychol. Monogr.*, 1935, 11, 1–42.
10. Lorge, I. Irrelevant rewards in animal learning. *J. comp. Psychol.*, 1936, 21, 105–128.
11. Roberts, W. H. The effect of delayed feeding on white rats in a problem cage. *J. genet. Psychol.*, 1930, 37, 35–58.
12. Schiller, P. H. Delayed response in the minnow. *J. comp. physiol. Psychol.*, 1948, 41, 233–238.
13. Schiller, P. H. Delayed detour response in the octopus. *J. comp. physiol. Psychol.*, 1949, 42, 220–225.
14. Schiller, P. H. Innate constituents of complex responses in primates. *Psychol. Rev.* (In preparation)
15. Schiller, P. H. Analysis of detour behavior: I. Learning of roundabout pathways in fish. *J. comp. physiol. Psychol.*, 1949, 42, 463–475.
16. Skinner, B. F. "Superstitions" in pigeons. *J. exp. Psychol.*, 1948, 38, 168–172.
17. Thorndike, E. L. *Animal intelligence.* New York: Macmillan, 1911.
18. Tolman, E. C., Ritchie, B. F., & Kalish, D. Studies in spatial learning, I–II. *J. exp. Psychol.*, 1946, 36, 13–24, 221–229.

LEARNING MOTIVATED BY A MANIPULATION DRIVE *

BY HARRY F. HARLOW, MARGARET KUENNE HARLOW, AND DONALD R. MEYER

University of Wisconsin

A preliminary study of puzzle solution in four rhesus monkeys suggested that a manipulation drive might operate in these animals and with sufficient strength to provide adequate motivation for learning. Psychologists have traditionally utilized the homeostatic drives in learning studies with subhuman animals, and have neglected, if not actually been blind to, the importance of externally elicited drives in learning.

The primary purpose of this study is to investigate the performance of monkeys on mechanical puzzles whose manipulation is accompanied by no extrinsic reward. A secondary purpose is to investigate the effect of the subsequent introduction of a food reward.

The study is comparable to the Blodgett (1) and the Tolman and Honzik (8) latent learning experiments, and the repetition of the Blodgett experiment by Reynolds (6), in that the initial training was conducted without special incentives, and later in training food incentives were introduced. The study differs from the earlier studies, however, in that the design of the present investigation places emphasis on the performance during the non-reward period, whereas the design of the other investigations places emphasis on the performance after the introduction of reward.

Subjects

Subjects in this investigation were eight rhesus monkeys with previous laboratory experience confined to discrimination, delayed reaction, and multiple discrimination reversal learning problems. The animals were matched in pairs and assigned to two groups on the basis of emotionality as judged by one person familiar with all the subjects under both living cage and experimental conditions. Group A, the experimental animals, contained Monkeys 138, 143, 150 and 151. Group B, the matched control subjects, included Monkeys 140, 146, 142 and 147.

Apparatus

The essential experimental apparatus, illustrated in Fig. 1, consisted of a metal-edged unpainted wooden base to which was attached, flush, a hasp restrained by a hook which was, in turn, restrained by a pin. The pin and hook are referred to as Restraining Devices 1 and 2, respectively, and had to be opened in serial order before the hasp, referred to as Device 3, could be raised. The apparatus was screwed to a wooden perch 24 in. long and 6 in. wide fastened to one side of an inside living cage 6 ft. long, 2½ ft. wide, and 6 ft. high. The animals at all times had access, through a runway, to an outside living cage 4 by 2½ by 6 ft.

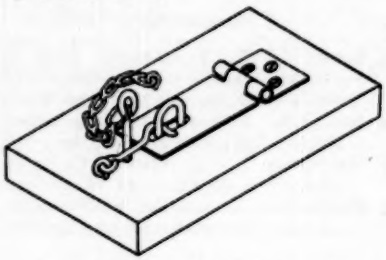


Fig. 1. Mechanical puzzle apparatus

For tests conducted in the experimental room, a puzzle identical with the one described except for a food-well one in. in diameter cut in the board and covered by the hasp, was screwed onto a 24 by 10 in. test tray placed on the floor of a 24 by 24 by 24 in. cage. A restraining cage 27 by 26 by 16½ in. adjoined the test cage.

Procedure

Throughout the experiment each subject was housed in individual quarters consisting of an

* Supported in part by the Research Committee of the Graduate School from funds supplied by the Wisconsin Alumni Research Foundation.

inside and outside cage, and fed daily, as is the customary procedure for the colony, between 5:30 and 6:00 P.M.

Tests without food reward.—During the first 10 days of the investigation, Group A subjects had the assembled puzzle attached to their perches and Group B subjects had the puzzle attached with the restraining devices unassembled and the hasp raised. At convenient times, an experimenter checked the puzzles in all the cages and recorded any successful manipulations that had occurred. Approximately 50 checks were made on each animal in the 10-day period. Any devices released by Group A subjects were reset at the time of the check. It was planned to unfasten any devices found assembled at the checking time for Group B subjects, but no instance of assembling behavior occurred. On both the 11th and 12th days, 10 checks were made on each subject at hourly intervals.

On both Days 13 and 14 five five-min. observations one hour apart were made on each subject. For the observation, the experimenter assembled the puzzle, left the cage, started the stop watch and observed from outside the cage without staring directly at the animal. A record was made in sec. of the time the subject initially touched the puzzle, released the restraining devices, and lifted the hasp. The order in which the parts were touched or manipulated was also recorded. At the conclusion of each of the observations except the last one on Day 14, the apparatus was reset if any devices had been released. Upon the experimenter's return an hour later or the following day for the next observation, a record was made of the successful manipulations accomplished in the interval.

Tests with and subsequent to food reward.—Approximately one hour before the daily feeding on Days 4 to 13 of the experiment, the animals in both groups were brought into the laboratory for adaptation to a special test situation to be administered on Day 14. The procedure on each adaptation day was to release the animal from the transport cage into the restraining cage, give him a few raisins, and then raise the cage door to permit entrance into the adjoining test cage. On the floor against the far side of the test cage rested a standard test tray 24 in. by 10 in., containing two one-in. food wells 12 in. apart, center to center. On the first five days, the subject obtained raisins from both food wells, took additional raisins introduced through the cage bars, then returned to the restraining cage. On each of the next five days, 20 delayed reaction trials were run. The monkey was kept in the restraining cage until the experimenter baited one food well and covered both wells with identical objects. The restraining cage door was then opened and the monkey allowed to displace the object and take the raisin. He then returned to the restraining cage and awaited the next trial. These tests were run merely to accustom the Group A animals to obtaining hidden raisins in the experimental cage situation. Group B animals were given identical training so that general adaptation would be constant for both groups.

The special puzzle test with food reward was conducted on Day 14 approximately two hours after the last living cage observation. Each Group A monkey was released into the restraining cage as on previous days. The tray in the test cage, however, now had the puzzle apparatus attached to it. While the animal watched, the food-well was baited with a raisin, the hasp lowered, and the restraining devices set. The restraining cage door was opened and the animal allowed to enter the test cage. A record was made of behavior for a five-min. period using the procedure described for Days 13 and 14 on the non-reward tests. The monkey was then returned to the living cage and immediately given a trial on his own puzzle apparatus which was baited in his presence with a raisin crushed beneath the hasp. A five-min. observation was made from outside the cage. The apparatus was reassembled immediately (this time without food) while the animal watched, and behavior was once more observed for five min.

Results

Tests without food reward.—The data of Table I present for each subject the total number of Day 13 and 14 solutions in 10 tests. A solution is defined as release of both restraining devices. The table also shows the number of solutions accomplished within 60 sec. of initial contact with the apparatus.

The Group A monkeys made a total of 31 solutions and the Group B monkeys a total of 4 solutions in 40 trials. The difference is significant at the one percent confidence level as determined by a *t*-test for related measures. There was no overlapping in the scores made by the individual animals of the two groups. None of the Group B monkeys ever solved the problem within 60 sec. of initial contact, whereas all the Group A monkeys solved the problem three or more times within this time limit. Over 50

TABLE I

PUZZLE SOLUTIONS ON DAYS 13 AND 14

Total Solutions				Solutions Attained Within 60 Sec.			
Group A		Group B		Group A		Group B	
Subject	Solutions	Solutions	Subject	Subject	Solutions	Solutions	Subject
138	8	1	140	138	5	0	140
143	6	1	146	143	3	0	146
150	9	2	142	150	7	0	142
151	8	0	147	151	7	0	147
Total	31	4			22	0	

percent of all solutions by the Group A monkeys were made in 20 sec. or less, and over 35 percent were made in 10 sec. or less. Monkey 151 twice failed to touch the puzzle in the five-min. test period but solved the problem without error in less than five sec. when the experimenter passed to the next cage.

The number of errors made before initial solution on Days 13 and/or 14 are given in Table II. An error is defined as contact with or manipulation of any device out of sequence, or any incorrect manipulation of the restraining devices. Three of the four Group A animals made the initial solution in the first observation period and without error. All Group B monkeys made errors before the initial solution and one monkey failed to solve the problem during the 10 five-min. tests.

TABLE II

NUMBER OF ERRORS BEFORE INITIAL SOLUTION ON DAYS 13 AND/OR 14

Group A		Group B	
Subject	Errors	Subject	Errors
138	0	140	11
143	6	146	16
150	0	142	7
151	0	147	15*

* Total errors. No solution attained.

The difference between the two groups as determined by a *t*-test of related measures, is significant at the one percent confidence level.

A comparison of the total number of correct and erroneous responses of the Group A and Group B monkeys is presented in Table III. All Group A monkeys made more than twice as many correct as incorrect responses, whereas all Group B monkeys made more than twice as many incorrect as correct responses. The differences are again significant at the one percent confidence level as measured by a χ^2 test of independence.

A further indication of the difference in proficiency of the two groups lies in the finding that the Group A monkeys made *no errors* in 21 of the 38 times that puzzle manipulations were observed, while the Group B monkeys *always made errors* in the 25 times that puzzle manipulations were observed.

Solution of the problem, it will be recalled, was defined in terms of releasing both Restraining Devices 1 and 2. Responses to Device 3 in the five-min. observation periods were variable (a finding confirmed in a subsequent study), even though all Group A monkeys invariably displaced the hasp in the one hour period between observations. Subject 138 ignored the hasp in one of eight solu-

TABLE III

TOTAL CORRECT AND INCORRECT RESPONSES ON DAYS 13 AND 14

Group A				Group B			
Subject	Correct Responses	Error Responses	Correct R / Incorrect R	Subject	Correct Responses	Error Responses	Correct R / Incorrect R
138	25	9	2.8	140	7	16	0.4
143	20	7	2.9	146	4	24	0.2
150	20	1	20.0	142	7	19	0.4
151	24	11	2.2	147	2	15	0.1

tions, Monkey 143 in two of six solutions, Monkey 150 in eight of nine solutions, and Monkey 151 in none of eight solutions.

Tests with and subsequent to food reward.—The data of Table IV show that in three of the four Group A subjects, the introduction of food rewards seriously disrupted the efficient puzzle solution which they had repeatedly demonstrated previously.

TABLE IV

ERRORS ON TESTS WITH AND SUBSEQUENT TO FOOD REWARD

F = failure; NR = no response

Subject	Test Room With Food	Home Cage With Food	Home Cage Without Food
138	5F	6F	3
143	NR	4	10
150	9F	1	1
151	0	2	0

In the experimental room test with food reward, Monkey 143 touched the apparatus and then retreated and refused to work, while Monkeys 138 and 150 worked on the problem, but failed after making 5 and 9 errors, respectively. On the food trial in the home cage, Subject 138 again failed, and the other three subjects made errors before their successful manipulations. It is to be noted that Monkey 143 performed very poorly in the subsequent home cage test without food.

The degree of disruption of the acquired puzzle responses in the case of Monkeys 138, 143, and 150 during and following the introduction of food, is indicated by their total of 39 errors in these three five-min. tests, although they had made a total of only 17 errors in the previous 10 five-min. tests. The difference in the *kind of errors* is also very striking. In the initial 10 tests without food reward, the monkeys *never* approached the problem by touching the hasp first; in the subsequent tests, all three monkeys *always* erred by attacking (literally) the hasp first. In the initial 10 tests without food reward the three monkeys had made a total of 2, 2, and 0 hasp errors, respectively, in contrast to the 8, 8, and 7 hasp errors made by the same monkeys subsequently.

DISCUSSION

The behavior obtained in this investigation poses some interesting questions for motivation theory, since significant learning was attained and efficient performance maintained without resort to special or extrinsic incentives. It is obvious, of course, that a number of physiological drives might have operated to activate the animals. Hunger must certainly have been present, since 14 to 22 hours had elapsed between the last feeding and the puzzle performances. Thirst may

have been present, at least part of the time, for water bottles occasionally run dry during the day. Data obtained in a subsequent study suggest that temperature drive played a role. Sex drive might also have operated. Granted that these physiological drives were present, the problem remains of explaining how such drives could operate to direct behavior on the puzzle problem. Solution did not lead to food, water, or sex gratification. The precise serial manipulations could not have provided greater temperature relief than random manipulation or locomotor behavior.

Whether one assumes a reinforcement or non-reinforcement theory of learning, the behavior of the monkeys is not explained by attributing their performance directly to the homeostatic drives. The reinforcement theorist is left to account for the drive or drives reduced by performance. The non-reinforcement theorist is faced with the problem of explaining the channeling of behavior.

The authors propose that an externally elicited drive operated to channel behavior on the mechanical puzzle, and that the performance of the task provided intrinsic reward. The idea of an externally-stimulated drive eliciting intrinsically-motivating behavior is not, of course, a new one. The exploratory-investigatory drive proposed by previous workers would fit in this category. The curiosity motive frequently attributed to human infants and children can also be characterized in this way. Tolman (7) has proposed play and aesthetic drives in apes and, presumably, human beings as primary drive conditions which might be externally elicited.

In spite of previous suggestions of such drives, little experimental work has been conducted relating to them, and the work that has been done appears to be limited to the exploratory drive. Nissen (5) obtained data on the behavior of rats in crossing a charged grid to a Dashiell checkerboard maze, and attributed the behavior to an exploratory drive. This same drive has been proposed by some to explain latent learning in the studies of Blodgett (1), Haney (3), Buxton (2), and others. In several of these experiments control over the possible escape incentive was obtained by housing the animals in the maze during the 'latent' learning part of the investigation. In another study, Mote and Finger (4) found a marked drop in latency of the simple running response of naive rats after an initial trial without food reward on the Graham and Gagne elevated runway. Because of lack of food reinforcement, however, the investigators did not interpret this improvement as a learning phenomenon.

An exploratory drive might have operated, in the present experiment, to initiate behavior towards the puzzle but some motivation in addition to the exploratory drive must be adduced to account for the persistence of the puzzle solving behavior through the course of the experiment without suggestion of extinction.

A secondary reinforcement explanation seems inadequate for a number of reasons. The only extrinsic incentive previously provided these subjects was food, and so presumably any secondary reinforcement would be derived from hunger reduction. Introduction of food in the present experiment served to disrupt performance, a phenomenon not reported in the literature on secondary reinforcement. Furthermore, the animals' history raises serious doubts about a secondary reinforcement interpretation. The previous experimental ex-

perience of the animals includes no problems of manipulation beyond pushing an object on a tray, and no problems of any kind presented in the living quarters. Some monkeys have been noted to reach through their living cages and manipulate the latches on the doors, but such responses have never led to solution and escape, and would thus provide no opportunity in the present experiment for secondary reinforcement of manipulatory responses based on escape. Additional evidence against a secondary reinforcement interpretation lies in the relative indifference during the observation tests of three of the four monkeys to the hasp, the only device that in any way resembled discrimination or delayed reaction objects. Finally, no evidence of extinction appeared after numerous trials without food or other extrinsic incentives. It seems far-fetched to propose that secondary reinforcement derived from hunger reduction in discrimination learning and delayed reaction problems conducted in the test rooms generalized to the puzzle problem presented in the living quarters, and once generalized, was maintained for 80 trials without any primary reinforcement.

Persons unfamiliar with monkeys might suspect social reinforcement of the responses, and theoretically it is possible that the experimenter might serve as a secondary reinforcing agent or provide reinforcement for a derived or secondary motive. The animals' history provides opportunities for such motivation, for taming procedures utilize food, and the human being is also associated with the daily feeding and watering. During the previous experimental work, however, interaction between subject and experimenter was minimized through the use of screens in all but one group of delayed reaction problems. General observation of the subjects has not suggested any affectional attachment to the experimenters which might have motivated performance. In the puzzle study, the experimenter was not present during the trials on the first 12 days, and during the observations on Days 13 and 14, the presence of the experimenter produced negative rather than positive effects. Certainly for Animals 146 and 151, the departure of the experimenter was a stronger stimulus than his presence to puzzle attack and solution. Social motivation is a possible explanation, but it appears to be an unlikely one.

It is the opinion of the experimenters that a manipulation drive can best account for the behavior obtained in this investigation. The stimuli to the drive are external and, in conjunction with the animals' capacities, set the pattern of behavior. The manipulation is conceived of as having reinforcing properties that account for the precision and speed the subjects acquire in carrying out the solution, and the persistence they show in repeated performances. Observations of subhuman animals and human beings provide abundant illustrations of directed manipulatory behavior unassociated with homeostatic drive reduction or social approval, but to the authors' knowledge no previous investigations have utilized this drive experimentally.

In conclusion we would like to emphasize the strength and the persistence of the manipulatory drive as described in this paper. It would appear that this drive, and probably a broad category of drives to which manipulation belongs, may be as basic and as strong as the homeostatic drives. Furthermore, there is some reason to believe that these drives can

be as efficient in facilitating learning as the homeostatic drives.

The homeostatic drives have been such convenient drives for the comparative-theoretical psychologist that the potentialities of other drives have been ignored and their role in behavior has been neglected or belittled. The unsatisfactory nature of modern drive theory may stem in large part from the biases of experimenters who have tended to cast all studies in a limited common mold.

Summary and Conclusions

1. Four rhesus monkeys were given 12 days' experience in manipulating a mechanical puzzle whose solution did not lead to any special incentive such as food or water. Four matched control subjects had the puzzles placed in their home cages the same period of time, but unassembled.

2. The performance of the two groups was then compared by observing their responses to the assembled puzzle during five five-min. periods on Days 13 and 14.

3. The experimental monkeys were more efficient than the control monkeys in puzzle solution as measured by total number of solutions, solutions attained in 60 sec., and ratio of correct to incorrect responses. The differences were significant at the one percent confidence level.

4. Subsequent introduction of food in the puzzle situation tended to disrupt, not facilitate, the learned performances of the experimental subjects.

5. A manipulation drive, strong and extremely persistent, is postulated to account for learning and maintenance of the puzzle performance. It is further postulated that drives of this class represent a form of motivation which may be as primary and as important as the homeostatic drives.

(Manuscript received April 11, 1949)

References

1. Blodgett, H. C. The effect of the introduction of reward upon the maze performance of rats. *Univ. Calif. Publ. Psychol.*, 1929, 4, 113–134.
2. Buxton, C. E. Latent learning and the goal gradient hypothesis. *Cont. to Psychol. Theory*, 1940, 2, No. 2, 3–75.
3. Haney, G. W. The effect of familiarity on the maze performance of albino rats. *Univ. Calif. Publ. Psychol.*, 1931, 4, 319–333.
4. Mote, F. A., & Finger, F. W. Exploratory drive and secondary reinforcement in the acquisition and extinction of a simple running response. *J. exp. Psychol.*, 1942, 31, 57–68.
5. Nissen, H. W. A study of exploratory behavior in the white rat by means of the obstruction method. *J. genet. Psychol.*, 1930, 37, 361–76.
6. Reynolds, B. A repetition of the Blodgett experiment on 'Latent Learning.' *J. exp. Psychol.*, 1945, 35, 504–516.
7. Tolman, E. C. Motivation, learning and adjustment. *Proc. Amer. phil. Soc.*, 1941, 84, 543–563.
8. Tolman, E. C., & Honzik, C. H. Introduction and removal of reward and maze performance in rats. *Univ. Calif. Publ. Psychol.*, 1930, 4, 257–275.

THE RELATIONSHIP BETWEEN TWO MEASURES OF RESPONSE STRENGTH

BY HAROLD SCHLOSBERG

Brown University

AND

CHARLES HEINEMAN

University of Iowa[1]

The strength of a conditioned response may be stated in terms of one or more of at least four different measures: amplitude, frequency, latency, and resistance to extinction. These measures have long been accepted on an empirical basis, without very much attention to the underlying mechanisms, or even to the inter-relationships among the various measures themselves. It is true that all of them have 'face validity,' and that they are inter-correlated to some extent (12), but until recently there had been little critical study of even one of these measures.

When Hull and his associates developed their systematic behavioral concepts, it was necessary to include reaction potential as an intervening variable. It was natural to take over the four measures from the field of conditioning (in the narrow sense) and to use them as indices of reaction potential. But in a system which has any claim to precision, it is necessary to examine the units which are used to measure the variables. Hence, Hull and his associates have devoted four papers (called the 'quantification series') (5, 9, 14, 22) to developing an equal-interval scale for measuring reaction potential in terms of latencies. But much remains to be done; we need to have some idea of the physiological mechanisms which make latency vary as some inverse function of reaction potential, and we should have formulae expressing the relationship of latency to other measures of reaction potential. Until these gaps are filled, reaction potential will be inadequately defined.

One of the difficulties that one encounters when he tries to fill this gap is the lack of adequate data. Large bodies of latencies are rarely available. Paired values for latency and some other measure of reaction potential are even more difficult to obtain. This is particularly true of latency vs. frequency, for in a simple (i.e., excluding discrimination) learning situation in which frequency of response is less than 100 percent, it is impossible to obtain a true cross-section of latencies. One cannot use even median latencies until learning has progressed to the point at which responses occur on at least 50 percent of the trials.

These difficulties can be overcome most readily in the familiar free association task, as the Kent-Rosanoff (15) type of experiment. One elicits a large number of responses, and measures the reaction time (latency) of each. He also has available the communality (frequency) of each response in the same data. It is true that communality is not a measure of the response of one individual, but for a homogeneous population this objection is probably not serious. Cer-

[1] The data on which this paper is based were collected and tabulated by the junior author as an Honors Project at Brown University. The senior author assumes responsibility for the theoretical treatment.

tainly the free association situation offers a relatively easy way to obtain large numbers of paired latency and frequency values for learned reactions. Any relationships found in this situation will be suggestive; they may then be tested in more conventional learning experiments.[2]

Thumb and Marbe seem to have been the first to investigate the relationship between association time and communality of response. Marbe's Law indicates that the more frequent responses are quicker. Since then the generalization has been amply confirmed (20). But unfortunately the data have not been presented in a fashion which permits more precise statement of the relationship. The early German workers (16, 21) used too few *Ss* to give stable results. Crane (1), who had 32 *Ss*, presents only mean association times, which do not adequately represent skewed distributions. Crossland (3) used 45 *Ss* in an attempt to pick guilty individuals. The raw data are presented in his monograph, but are not worked up in a fashion convenient for present purposes. Further, they are stopwatch times, obtained in an emotion-laden situation. Hence, it seemed more satisfactory to collect new data.

The present study is concerned with the distribution and interrelations of association time and communality values obtained from responses of 204 undergraduates to 25 monosyllabic words from the Kent-Rosanoff list.

Procedure

An unselected group of 204 pairs of students of elementary psychology at Brown University was required to respond to each of the 100 stimulus words of the Kent-Rosanoff list (15). There were about twice as many men as women in the group, but the self-chosen partners tended to be of the same sex. One member of each pair acted as *E* and the other as *S* for the first 50 words. The responses and reaction times used in this report were obtained from the 204 students who acted as first *S*. They were, therefore, naive in regard to the words. *E* was asked to note any unusual behavior, nervousness, or signs of embarrassment on the part of *S*.

The *Es* recorded the moment of stimulation and response by actuating a toggle switch on a reaction apparatus containing a 1/100 sec. Springfield Timer (17). The raw data were later rounded to $\frac{1}{10}$ sec. for statistical work.

Several years before the present study, the senior author had directed a pilot study, utilizing several practices *Es* and 50 *Ss*. On preliminary analysis of the data it was clear that a larger number of *Ss* would be needed to give stable results. This suggested the use of students in the laboratory sections of the large elementary course (17). Careful consideration was given to the merits of using 200 unpracticed *Es* in comparison with one practiced *E* and fewer *Ss*. Objective conditions would be controlled better by a practiced *E*, but the extremely important subjective factors, as attitude and emotion, might well be more uniform when *E* and *S* were laboratory partners. These pairs had worked together, each pair in its own enclosed booth, for a half-dozen laboratory periods before this experiment. Further, they had already used the timers in two earlier experiments, so they were familiar with their operation. Finally, unusual care was taken to ensure that the students performed the experiment properly. The directions had been refined over a period of several years.[3] The students were told that they were participating in research, and not merely performing a routine experiment. As an extra precaution, the laboratory assistants made careful checks to see that the directions were being followed.

The other major question that arose was how to time the reactions. Voice keys were rejected, not only because they are cumbersome, but chiefly because some *Ss* form the habit of stopping the clock by a preliminary 'Aaah' before the response word. Similarly, many *Ss* cannot be depended upon to flip a hand switch in coincidence with the associative response. On the other hand, *E* can snap a switch as he says the word, and snap it back when *S* responds. This method may introduce a relatively constant error, but it is important only for short reactions.

[2] In the free association experiment the construct being measured, associative strength, includes motivating as well as learning factors. It is thus analogous to response strength, reaction potential (13) or strength of the conditioned response, in contrast with strength of conditioning (12) which is comparable to habit strength (13).

[3] Copies of the directions, and of frequency tables, will be furnished on request.

A closely related question concerns the exact moment of stimulation. Should *E* start the timer on the first syllable, on the accented one, or on the final one in a long word? In the present experiment *E* was instructed to start the timer on the stressed syllable of the word. However, one could argue that this method arbitrarily adds a half-second to a word like 'Butterfly.' The easiest way to minimize this type of error is to eliminate all long stimulus words. Hence, the present paper will consider only the 25 monosyllable words which appear in the first half of the Kent-Rosanoff list.

Results

The primary data of the present experiment are the responses and their latencies elicited by 25 monosyllabic words, from each of 204 naive *S*s. Only 18 out of the possible 5100 responses were lost by failure of *S* to respond, or of *E* to record them legibly. In Table I these data have been grouped by communality, and listed with corresponding median reaction times. It will be seen that there were 630 unique or individual responses, and that their median reaction time was 1.74 sec. There were 202 cases in which the same S-R combination occurred only one other time; that is, had a communality of 2. The median reaction time for these responses was also 1.74 sec. But as one continues down the list, he will notice that median reaction time decreases as communality increases; at the bottom of the list a very frequent combination (it was 'Slow—Fast') occurred 171 times out of a possible 204, or 84 percent, and had a median reaction time of 1.1 sec. This inverse relationship between communality and reaction time is what one expects, but before more can be said it will be necessary to examine the distributions of the separate reaction times.

TABLE I

The Incidence and Median Latency of Associative Responses of Various Degrees of Communality. The communalities are based on a population of 204 students; hence they may be converted into rough percentages by halving them.

Communality	Incidence	Reaction Time (in sec.)	Communality	Incidence	Reaction Time (in sec.)
1	630	1.74	29	29	1.5
2	202	1.74	30	30	1.34
3	123	1.54	31	31	1.44
4	84	1.5	32	64	1.41
5	100	1.66	36	36	1.6
6	84	1.5	40	40	1.5
7	42	1.45	55	55	1.14
8	56	1.45	59	59	1.4
9	63	1.4	65	65	1.17
10	10	1.35	77	231	1.4
11	33	1.53	78	78	1.4
12	48	1.4	86	86	1.23
13	39	1.5	87	87	1.23
14	84	1.54	93	93	1.3
15	30	1.3	119	119	1.2
16	16	1.95	122	122	1.29
17	51	1.6	126	126	1.34
18	36	1.5	132	132	1.2
19	76	1.4	133	133	1.18
20	20	1.4	138	138	1.1
21	63	1.4	147	147	1.22
22	22	1.5	156	312	1.1
23	23	1.43	159	159	1.1
24	48	1.5	161	161	1.26
25	25	1.4	165	330	1.08
28	28	1.54	171	171	1.1

The distribution of association times

It is common knowledge that distributions of latencies are usually skewed. Since most statistical techniques are based on a normal distribution, it is convenient to scale raw latent times to normalize distribution. A logarithmic scale has often been used for this purpose (6, 10, 19). In addition to the practical advantage that most statistical techniques assume normality, there is a theoretical point involved. Hull (13) first assumed that reaction potential (${}_{S}E_{R}$) is normally distributed.[4] This was one of the basic assumptions on which

[4] But in a later paper, it was shown that reaction potential is distributed in a leptokurtic fashion.

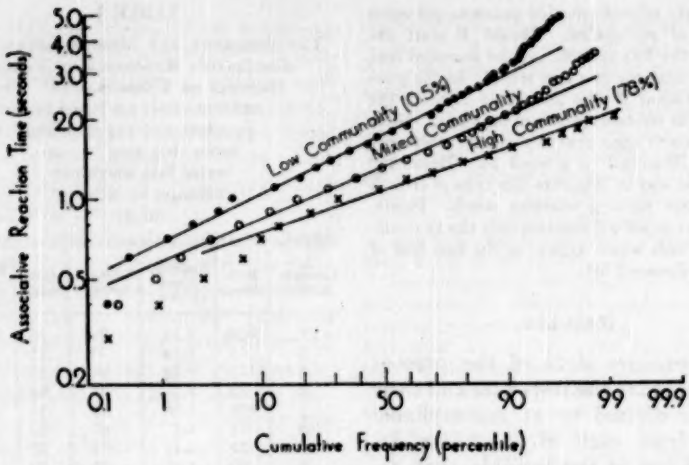


FIG. 1. The distribution of associative reaction times for three degrees of communality. The ordinate is logarithmic, and the abscissa is spaced to yield a straight line for a cumulative plotting of a normally distributed variable. The hollow circles represent 1000 reactions of mixed communality. The solid dots are based on 630 unique responses, and the crosses on 312 responses with a communality of 156.

the quantification of reaction potential was based. So far as the assumption is correct, a relatively simple treatment of reaction time values, designed to normalize their distribution, should also give a fair scaling of reaction potential (or associative strength, in the present experiment).

An easy way to test the approximate normality of a distribution is to plot it on a probability grid. This graph paper has the base line so scaled in percentile units that it yields a straight line when a normally distributed variable is plotted as cumulative frequency. The paper is obtainable with either arithmetic or logarithmic ordinates.

All reactions of the first 40 *Ss* were tabulated, and the 1000 reaction times thus obtained were plotted as a set of cumulative frequencies, using the logarithmic probability paper. These values are shown as hollow circles in Fig. 1. The straight line drawn through these points gives a fair fit from the 4th to the 94th percentile, corresponding to association times of .7 to 2.4 sec. It is clear that the distribution of log association times is roughly normal, but that there is a slight expansion at the ends of the distribution. The curve as a whole suggests a slight degree of leptokurtosis, somewhat less than that indicated by a replotting of data from Hull *et al.* (22, p. 225) for a distribution of latencies. But detailed analysis of these 1000 reaction times is unprofitable, for they are based on a heterogeneous group of reactions, varying in communality from 0.5 to 84 percent. Instead, it seems more desirable to examine several latency distributions, each obtained from responses of a given degree of communality.

In Fig. 1, the dots represent cumulative frequencies of association times for 630 unique reactions (communality of 0.5 percent), while the crosses represent similar data for 312 responses with a communality of 76 per-

cent. These are extreme degrees of communality. The straight lines drawn through these points again yield reasonable fits, except at the extremes. The departures from normality, however, are instructive. The upper 10 percent of the unique responses are too slow, while the lower 10 percent of the common responses are too fast. Similar distributions for seven intermediate degrees of communality show a gradual shift from one type of distortion to the other (Table II). The breakdown points are more constant if expressed in time values, rather than in percentiles. It is a combination of these distortions that yields the distribution found for 1000 responses of mixed communality. Hence, one should not devote further attention to the heterogeneous distribution, for its exact form will depend on the original choice of stimulus words.

Several attempts were made to find a simple conversion that would yield more nearly normal distributions of latencies. The addition or substraction of a constant of the order of 0.5 sec. to latencies before conversion to logs, corrects one end of the distribution at the expense of the other. Simple reciprocals of reaction time are normally distributed only if a relatively large constant is added before inversion. A negative fractional power function ($RT^{-0.5}$) yields distributions that are reasonably normal from the 10th or 15th to the 99th percentile, but the short reactions are much too short. This is the type of function which emerged as a result of the Hull quantification series. These various scales will be considered in the Discussion. For the present, it is sufficient to indicate that log association times give a reasonably normal scale.

TABLE II

THE DISTRIBUTION OF ASSOCIATIVE REACTION TIMES FOR RESPONSES OF DIFFERENT DEGREES OF COMMUNALITY

Communality (out of 204)	Linear Part of Range		Medians		SD (Log)
	Percentiles	Seconds	(Sec.)	(Log)	
1	.4–90.0	.6–3.0	1.74	.24	.18
2	.5–92.0	.6–3.1	1.74	.24	.18
3	6.0–92.0	.8–2.8	1.54	.19	.18
5	9.0–97.0	1.0–3.3	1.66	.22	.15
14	1.0–91.0	.9–2.5	1.54	.19	.15
32	2.0–95.0	.8–2.3	1.41	.15	.13
77	7.0–91.0	.8–2.6	1.40	.15	.14
122	7.0–99.0	.8–2.6	1.29	.11	.13
156	12.0–99.0	.8–2.2	1.10	.04	.13
Varied*	4.0–94.0	.7–2.4	1.30	.11	.15

* Varied includes all 25 responses from each of 40 Ss.

The distribution of communality, and its relation to latency

Some effort was devoted to finding a simple conversion factor that would yield a normal distribution of communalities. However, it was soon apparent that the attempt was futile, for the distribution was roughly U-shaped, representing a combination of two J-curves. If one considers the number of different responses with a communality of 1, 2 . . . N, it is obvious that he is dealing with a curve that falls rapidly. On the other hand, the actual number of times that the same word appears is a linear function of communality, by definition. The net result is that the frequency is highest for unique responses, falls steeply and reaches a minimum around 15 percent communality, and then climbs steadily as communality increases. Such distributions are often lawful (18), but they must also be influenced by the constitution of the list.

Since communality does not readily yield a scale by direct treatment, some other method must be employed.

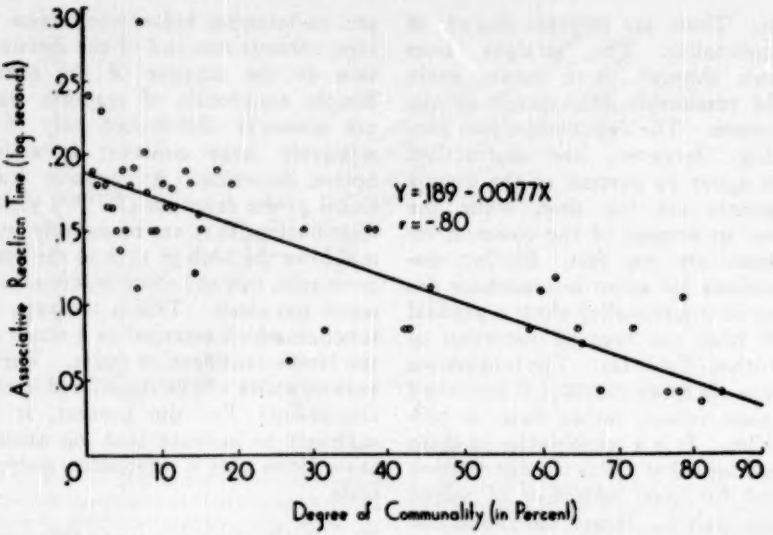


FIG. 2. The relationship between associative reaction time and degree of communality. Each point represents the median log reaction time of all responses which occurred with a given degree of communality. In the equation, Y is in log seconds, and X is in percent. (Data from Table I.)

Fortunately, communality can be related to median latency. That is, if one plots each communality against median association time of all responses of that degree of communality, some function should emerge. If association time is already scaled in a fairly satisfactory way, communality can be scaled by conversion.

In Fig. 2, communality, in percent, has been plotted as a function of median latency. The ordinate is logarithmic, since it has been shown above that this function yields a reasonable latency scale. The straight line drawn through the values represents the equation $\log RT = .189 - .00177$ communality. Its adequacy may be judged by a correlation of $-.80$ between log association time and percent communality.

The fit of this line is far from ideal, but no better one could be found by relatively simple manipulation of ordinate and abscissa. The various conversions of association time mentioned in the previous section have very little effect, for the range of median times is small. Slight improvement can be obtained by using the square root of communality, which brings the first two (low communality) time values nearer to the line, but the total improvement is negligible. It is probable that some more complex function (e.g., a Gompertz curve) would give a better description of the relationship, but the stability of the data would scarcely support any constants that might be derived in this fashion. Hence, it might be well to be satisfied with the generalization that log association time is inversely related to communality in a roughly linear manner. This relationship is probably precise enough for practical purposes; it will be shown in the Discussion that the relationship has some rational basis.

Results from the pilot study

It will be recalled that a pilot experiment was mentioned in the Procedure. The results of the 50 *Ss* who participated in that experiment were examined again in the light of the data obtained in the present experiment. Distributions of log association times from the pilot study were similar in form to those obtained from the larger group if allowance is made for the irregularity incident upon a small sample, and the inclusion of polysyllabic stimulus words. Similarly, the plot of log latency against communality had the same general form in both studies. There was a tendency for the slope to be greater in the pilot study, but the same y-intercept would serve for both sets of data.

There is one additional bit of information that may be furnished by this study. The distributions of latency vs. communality had been made separately for each quarter of the list. The median latencies for most communalities tended to fall as *S* progressed through the list. In other words, practice or familiarity tended to decrease association time (cf. 20). The data of the pilot experiment are not stable enough to permit a methematical statement of this change.[5]

Discussion

A. The empirical relationships among association time, communality, and strength of association

1. *Practical considerations.*—The present results indicate that one may legitimately perform many statistical operations on association times if they are first normalized by converting them into logarithms. This conclusion is not unexpected, and may well hold for most latency measures (6, 10, 19). But far too many workers, both in psychology and in other fields, proceed to add and average values from badly skewed distributions. The resulting means do not adequately describe the data, even if they are accompanied by measures of dispersion. The situation would not be as bad as it is if these writers followed the old practice of giving medians as well as means, for the median at least represents a typical score and can be converted into logarithmic or other scales if desired. However, if only the means are available in a published paper, further statistical manipulation is limited to the scale which was tacitly assumed in the original computations. Thus, if one tries to examine the relationship between association time and communality from Crane's data (1) as Woodworth has done (20), he is forced to use an arithmetic ordinate.

2. *The relationship between association time and strength of association.*—The next question to be raised is whether or not log association time is a linear measure of strength of association or, more generally, of reaction potential (sE_R). If one follows the reasoning on which the Hull quantification series (5, 9, 14, 22) is based, the answer would be, "Approximately so." That is, latent time (and presumably the sub-class, association time) is by definition a measure of reaction potential. Further, reaction potentials are assumed to be approximately normally distributed (13) and log association time is similarly distributed; hence, the relationship between the two might well be linear. Finally, either the logarithmic function, or the negative power function which was found to fit the present data almost as well, conforms surprisingly well to those which resulted from the elaborate treatment developed by Hull and his associates. On empirical grounds, there is little basis to choose one of these functions rather than the other. The logarithmic conversion exaggerates the very short and very long reactions, whereas the negative fractional

[5] The writers wish to thank Mrs. Marion Dahms Holland, who made tabulations and preliminary plots of these data. The pilot experiment was of great assistance in planning the main experiment.

power function concentrates all the distortion on the short-reaction end of the distribution. Either gives a satisfactory fit for the middle 80–90 percent of association times. It is possible that the distortions at the short end are due to errors in measurement; a few cases in which *E* gave the stimulus word before he snapped the switch would show up clearly in short reactions. Distortions at the other end might be attributed to the introduction of a new factor, as embarrassment, when *S* doesn't respond within two or three sec. Hence, it may not be desirable to try to include extreme times in a single function.

Distortions of this type are not limited to the present experiment. Hull and his associates (22) found that their original formula for reaction time yielded a distribution of latencies that showed some kurtosis. Furthermore, the distribution of momentary reaction potential, based on corrected latencies at the asymptote of learning (22, Fig. 1), is slightly asymmetrical. Values derived from this figure, when plotted on probability paper, show about as much distortion as do the log association time plots reported here.

If one wishes to make the whole range of association times conform to a normal distribution, it can be done by adding a constant, of the order of one sec., to the raw times, and then converting the sums to reciprocals. But this is empirical curve-fitting at its worst, for it will be shown below that there are no rational grounds for the addition of such a constant. Hence, it would seem preferable to use the logarithmic conversion, on the grounds of simplicity.

3. *The relationship between association time and communality.*—A second line of reasoning also suggests that log association time is a linear measure of strength of association. It has been shown above that log association time varies inversely with communality. The latter measure is a special case of frequency of occurrence, which is also widely used as a measure of reaction potential. It is rare in the field of learning to find that two supposed measures of the process correlate to the extent of .80. To the writer's knowledge, the highest correlation that has previously been reported between latent time and frequency of occurrence was −.54 (12). The high correlation found in the present study undoubtedly reflects the high reliability which was obtained by using median values based on a large number of cases. But the fact that the correlation is as high as it is furnishes additional evidence that log association time and communality are both measuring the same variable. But what is this variable? It would be simple to call it 'reaction potential,' thereby classifying it with a familiar construct, and avoiding further trouble. However, it might be instructive to examine it further, to see what may reasonably be happening during the association process.

B. *Rational approaches to association time*

1. *The result of oscillation.*—There are two obvious ways in which reaction potential (or associative strength) could be converted into latent time (or associative reaction time). The first of these utilizes the assumption that there is a rapid and random oscillation of reaction potential [6] (14, p. 240). Whenever momentary reaction potential exceeds the reaction threshold, the response will occur. If average reaction potential associated with a specific stimulus is weak, there will tend to be a long interval (latent time) before an extreme oscillation brings momentary reaction potential above threshold. Conversely, if a stimulus releases a strong reaction potential, whose oscillations are largely above threshold, the latent time will be short. This line of reasoning fits in well with Hull's general approach.[7]

[6] It is possible that this oscillation should be referred to the reaction threshold. Even if both reaction threshold and reaction potential show normally distributed and independent variations, the argument still applies, for the differences between them will still be normally distributed.

[7] In *Principles of behavior* (13, pp. 325–335) Hull ingeniously derives reaction evocation probability through the effect of oscillations on re-

If one assumes normally-distributed oscillations of reaction threshold as the basis of latent time, it would seem to follow that the latent times would yield a J-shaped distribution, or decay function, whose slope would depend on the strength of average reaction potential. This can be demonstrated by considering distributions of latencies at three stages of learning, chosen so that mean reaction potential is (a) equal to reaction threshold, (b) well above threshold, and (c) well below threshold. In each case it will be assumed that the oscillations occur with a frequency of one per sec.,[8] and that the amplitudes of successive oscillations are random with respect to preceding ones. Each of the three distributions will be based on 100 cases: one trial from each of 100 animals at the same level of learning and drive.

a. If the mean reaction potential is equal to the reaction threshold, there is a 50 percent chance that the first oscillation will bring momentary reaction potential above reaction threshold. Thus, 50 of the 100 responses will occur during the first sec. There will still be 50 responses that are yet to occur; 50 percent of these remaining responses, or 25, will fall during the next sec. Continuing the series, successive seconds will contain 13, 6, 3, . . . etc., responses.

b. Assume that mean reaction potential is so far below reaction threshold that the probability of an oscillation carrying momentary reaction potential above threshold is 10 percent. Following the method of applying this percentage only to the responses which have not occurred before the beginning of a given second, the distribution of latencies on successive seconds will be 10, 9, 8, 7+, 7−, 6, 5+, 5−, 4, etc., tailing off gradually over a long range of time values.

c. If the mean reaction potential is so far above reaction threshold that the probability of response during the first second is 90 percent, one would expect 9 percent of the responses during the next sec., 0.9 percent during the third, and the remaining 0.1 percent during the next two or three sec.

The J-shaped distributions predicted by this hypothesis are clearly *not* in agreement with empirical results. Hence, the hypothesis must be discarded.

2. *A functional analysis.*—The basic trouble with the above hypothesis is that it assumes reaction potential to spring into being at full strength (except for random oscillations), as soon as it is elicited by the stimulus. Such an assumption is inconsistent with what is known about physiological processes in general, and potentials in particular (8). Perhaps the most relevant material comes from the extensive work of Eccles (see 4 for references), who has recorded the gradual rise of electrical potentials in various synaptic regions of the sympathetic and central nervous system. A single nerve impulse will cause the potential to build up to an asymptote, and then to fall off slowly. If a train of impulses falls on the area, the electrical potentials summate, giving a gradual build-up to a higher asymptote. In the spinal cord of the frog, the build-up may take .15 sec., but in the stellate ganglion it may last as long as .4 sec. When such potentials are measured in the cerebral cortex, it is possible that they will be considerably slower. Eccles believes that these electrical potentials represent a critical step in synaptic transmission (4). As he conceives the process, when the potential reaches a certain

action potential and reaction threshold. He follows with two pages on reaction latency, but makes no effort to carry out a similar analysis of this measure.

[8] The examples are intended to give the general form of distributions predicted by this hypothesis. The form would not be changed by assuming that the oscillations were more or less rapid than one per sec. It should also be noted that the analysis is concerned with reduced latent times; peripheral factors like receptor and effector delay, as well as possible apparatus constants, might change the basic curves if raw latencies were used. (See Discussion, Sect. B, 2.)

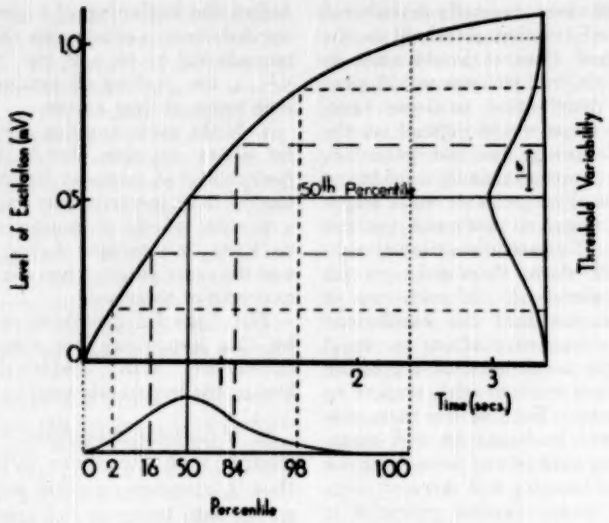


FIG. 3. A physiological model which predicts the general form of the distribution of associative reaction times.

The negatively accelerated curve represents the development of potential as a function of the time since onset of the stimulus. It is a tracing of one obtained by Eccles (3) from the ninth ventral root of the frog's spinal cord, set up by continued stimulation of the ninth dorsal root. Eccles' ordinate scale (mV) has been retained, but the time values on the baseline have been multiplied by a factor of 20.

A normal distribution of threshold has been indicated on the right hand ordinate. Horizontal lines, spaced by sigma distances of this distribution, have been drawn to intercept the curve of developing potential. The time values of these intercepts have been used in constructing the distribution of latencies drawn below the figure. The predicted distribution is markedly skewed, as are those generally obtained from latency measurements.

level (the threshold), the next unit in the chain (the post-synaptic fiber) discharges.

This same general picture seems to hold for peripheral nerve, neuro-muscular junction, sympathetic ganglia, and spinal cord. In each case there is a constant input of stimulating energy, which builds up a potential toward an asymptote. When the potential passes a threshold value, the next unit discharges. We are thus furnished with a model for any S-R sequence. In terms of the association experiment, the stimulus word releases excitation which builds up toward an asymptote at some 'center.'[9] When the level of excitation reaches a threshold, the center will discharge, and the response will occur. If one assumes that the threshold varies in a normally distributed manner, it is possible to predict the distribution of associative reaction times. In Fig. 3 the curve is a tracing of a record of the development of electrical potential in the ninth ventral root of the frog's spinal cord, resulting from stimulation of the ninth dorsal root at a rate of

[9] The word 'center' is used in a functional sense. The reader may substitute 'mechanism' or 'region' if he wishes.

200 per sec., for .16 sec. (3). The baseline has been multiplied by a factor of 20 to make the times more nearly comparable with those of the association experiment. The horizontal lines drawn across the figure represent excitation levels required to cross threshold as the latter varies in a random fashion. They originate from the assumed distribution of thresholds indicated on the right hand ordinate. Thus, the time value at which the solid line (labelled 50%ile) crosses the curve represents the median latency. Similarly, the horizontal broken lines represent thresholds that are spaced by one or more sigma above and below the mean, and permit the determination of latencies in terms of expected percentiles. From these values, with the aid of a table giving the height of the ordinate at various sigma distances in a normal curve, it is possible to construct the expected distribution of latencies drawn below the figure. These distributions show the same skew that was found in the actual associative reaction times. As a matter of fact, if one subtracts 0.5 sec. from the raw reaction times for responses with a degree of communality of 60 percent, thereby reducing the reaction times to 'central' times (see below), he will obtain a distribution with approximately the same absolute characteristics as that derived in Fig. 3. The absolute correspondence was the chance result of fortunate choice of curve and constants, rather than a deliberate attempt at curve-fitting; the general similarity between the derived and empirical distributions, on the other hand, may be taken as an argument for this type of model.

The above model is obviously much too simple. In the first place, the curve representing the rise in potential is only one of a family of possible curves, corresponding to differences in strength of the stimulus, and in the efficiency of the receptor and conduction paths between receptor and center. Undoubtedly habit strength would be included under this heading. The detailed form of the curve is not known, even for the spinal level. Further, no attention has been paid to inhibitory processes that would be released by the stimulus word, thus reducing the rate of accumulation of excitation. Finally, the threshold itself may involve at least two groups of factors. The first of these is the threshold proper, or the level of excitation which must be reached before it discharges into response. The chemical aspects of drive may influence this level.[10] The second group of factors subsumed under threshold is the background level of excitation of the center before the energies released by the immediate stimulus start to build up. The general tonic level of the body would seem to operate through a mechanism such as this (7). Similarly, the effect of the phenomena called 'set' might affect the background level of specific centers. The stimulus aspects of drive would probably work in the same fashion.

Any or all of the factors listed above may vary in a more or less random fashion, with consequent variations in 'central' time. But the raw latent times also include some additional peripheral components. There are apparatus lags, sense organ delays, muscle inertias, and conduction times to and from the central nervous system. For example, the distributions of latencies for conditioned eyelid and knee jerk would differ if for no other

[10] There is considerable evidence that direct chemical effects on the nervous system are of importance in certain drives, as hunger, thirst, and sex. See Morgan, C. T., *Physiological psychology*.

reason than the relative mass of the responding members.

One could summarize this part of the discussion by writing a long equation, with a constant for each of the factors mentioned above. The numerical determination of these constants would not be complete in one lifetime, even for a single type of reaction. But there is reason to believe that many of the factor involved would vary as decreasing functions, as do the increments of excitation shown in Fig. 3. These functions seem to be roughly logarithmic through much of their range. Hence when they combine, they yield a total latent time that is normally distributed if converted to logarithms. The situation is quite analogous to that of Weber's Law. This law is only a rough empirical generalization. Finer analyses of the processes involved in vision, for example (11), are resulting in a rational analysis and more precise statement of the functions underlying the general law, but they do not destroy its practical utility. In the same way, it is useful to know that log association time is distributed in a roughly normal fashion, particularly in the mid-ranges of time values.

As a byproduct of this analysis, the significance of communality is reasonably clear. Each stimulus word releases excitation to 'centers' for many different response words. The first 'center' in which net excitation reaches a threshold value will discharge, and the corresponding response word will occur. Thus, communality (and percent frequency, in general) represents the number of times a particular response wins the race. All the factors which determine association time will influence this competition. Hence, communality will vary inversely with association time. The exact form of this relationship will depend on the competition a specific response meets with from other responses associated with the same stimulus word. For practical purposes, the relationship may be taken as linear.[11]

Summary and Conclusions

1. Associative responses and their reaction times were recorded for each of 204 undergraduates, using 25 monosyllabic stimulus words. Each of the 5000 responses could be described in terms of two measures of associative strength: (1) associative reaction time (latency), and (2) communality (percentage frequency).

2. Associative reaction times, after conversion to a logarithmic scale, were distributed in an approximately normal fashion. Times below 0.9 sec. tended to be too short, while those above approximately 2.5 sec. were too long. A scale based on a negative fractional power function was slightly less satisfactory.

3. When reactions were segregated in terms of communality, the median log reaction time was found to be an inverse linear function of communality. The correlation between log latent time and percentage frequency was −.80.

4. An analysis of the association process in terms of an oscillating associative strength and normally distributed thresholds predicted a J-shaped distribution of reaction times. On the other hand, an analysis based on the rate of development of stimulus-released energies predicted distributions of latency which were consistent with those obtained from the experiment. A further consideration of possible physiological factors

[11] The senior author wishes to thank Drs. Pfaffmann, Kimble, and Zeaman for their many helpful suggestions and critical comments.

which are involved in the associative reactions indicates that we will have to be satisfied with a rough empirical function, which resembles Weber's Law in that its constants will differ from situation to situation, and that it will hold only for the mid-ranges of the scale.

(Manuscript received May 23, 1949)

References

1. Crane, H. W. A study in association reaction and reaction time. *Psychol. Monogr.*, 1915, 18, No. 80.
2. Crossland, H. R. The psychological methods of word association and reaction time as tests of deception. *Univ. Ore. Publ. Psychol.*, 1929, No. 1.
3. Eccles, J. C. Acetylcholine and synaptic transmission in the spinal cord. *J. Neurophysiol.*, 1947, 10, 197–204.
4. Eccles, J. C. An electrical hypothesis of synaptic and neuro-muscular transmission. *Ann. N.Y. Acad. Sci.*, 1946–7, 47, 429–455.
5. Felsinger, J. M., Gladstone, A. I., Yamaguchi, H. G., & Hull, C. L. Reaction latency ($_{sE_R}$) as a function of the number of reinforcements. *J. exp. Psychol.*, 1947, 37, 214–228.
6. Finger, F. W. Quantitative studies of "conflict." I. Variations in latency and strength of the rat's response in a discrimination-jumping situation. *J. comp. Psychol.*, 1941, 31, 97–127.
7. Fulton, J. F. *Physiology of the nervous system.* N. Y.: Oxford Univ. Press., 1938.
8. Fulton, J. F. (ed.) *Howell's textbook of physiology.* Philadelphia: Saunders, 1947.
9. Gladstone, A. I., Yamaguchi, H. G., Hull, C. L., & Felsinger, J. M. Some functional relationships of reaction potential ($_{sE_R}$) and related phenomena. *J. exp. Psychol.*, 1947, 37, 510–526.
10. Graham, C. H., & Gagné, R. N. The acquisition, extinction, and spontaneous recovery of a conditioned operant response. *J. exp. Psychol.*, 1942, 26, 267–288.
11. Hecht, S. Vision: II. The nature of the photoreceptor process. Chapter 14 in: Murchison, C. *Handbook of general experimental psychology.* Worcester: Clark Univ. Press, 1934.
12. Hilgard, E. R., & Marquis, D. G. *Conditioning and learning.* New York: Appleton-Century, 1940.
13. Hull, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.
14. Hull, C. L., Felsinger, J. M., Gladstone, A. I., & Yamaguchi, H. G. A proposed quantification of habit strength. *Psychol. Rev.*, 1947, 54, 237–254.
15. Kent, G., & Rosanoff, J. J. A study of association in insanity. *Amer. J. Insan.*, 1910, 67, 37–96.
16. Menzerath, P. Die Bedeutung der sprachlichen geläufkeit oder der formalen sprachlichen Beziehung für die Reproduktion. *Z. Psychol.*, 1908, 48, 1–95.
17. Schlosberg, H. A laboratory period in the first course in psychology. *Amer. Psychologist*, 1947, 2, 384–387.
18. Skinner, B. F. The distribution of associated words. *Psychol. Record*, 1937, 1, 70–76.
19. Solomon, R. L. Latency of response as a measure of learning in a "single-door" discrimination. *Amer. J. Psychol.*, 1943, 56, 422–432.
20. Woodworth, R. S. *Experimental psychology.* New York: Henry Holt, 1938.
21. Wreschner, A. Die Reproduktion und Assoziation von Vorstellungen. *Z. Psychol.*, Ergänzbd. 3, 1907–9.
22. Yamaguchi, H. G., Hull, C. L., Felsinger, J. M., & Gladstone, A. I. Characteristics of dispersions based on the pooled momentary reaction potentials ($_{s\bar{E}_R}$) of a group. *Psychol. Rev.*, 1948, 55, 216–238.

EVIDENCE FOR THE ROLE OF MOTIVATION IN DETERMINING THE AMOUNT OF REMINISCENCE IN PURSUIT ROTOR LEARNING

BY GREGORY A. KIMBLE

Brown University

Introduction

The results obtained in this investigation are presented as partial confirmation of two hypotheses developed more fully in earlier publications (2, 3). The first of these is that the amount of reactive inhibition (I_R) which an S will tolerate in a motor learning situation is partly a function of his motivation to do well at the task he is attempting to learn or perperform: the greater the amount of this motivation, the greater the amount of I_R he will tolerate. This hypothesis follows from the fact that reactive inhibition is thought of as a negative drive, i.e., as need to rest. As such, it acts in opposition to the needs leading to the continuation of practice. If we adopt the position that a given amount of I_R cancels a certain amount of these more positive needs, then it follows that a highly motivated S will continue to practice in spite of the presence of an amount of inhibition which would produce a cessation of activity in a less highly motivated S. The experimental prediction to which this argument leads is that the amount of reminiscence, which is correlated positively (perhaps perfectly) with the amount of I_R, will increase with an increase in motivation.

The second hypothesis is that conditioned inhibition ($_SI_R$) will not develop in the motor learning situation except under conditions of extreme massing which are continued for several minutes. This prediction follows from the postulation of a critical level or threshold value of reactive inhibition which must be reached and maintained before conditioned inhibition can begin to develop. If practice sessions are sufficiently short and if the inter-session rests are long enough to allow I_R to dissipate almost completely, no $_SI_R$ should develop because the threshold value of I_R will never be reached. The experimental test of this hypothesis involves the demonstration that a work-rest sequence can be found such that during the practice session performance is depressed; but that after a rest initial performance is equal to that of a group of Ss working under near optimal practice conditions. Since $_SI_R$ is considered as producing a permanent decrement in performance, such a demonstration would indicate that I_R but no $_SI_R$ had developed under these particular learning conditions.

Experimental Design and Procedure

The apparatus used in this investigation was a modified Koerth pursuit rotor which has been described in its essential details in an earlier paper (4). For purposes of the present experiment, a Springfield Electric Timer was substituted for the impulse counter previously used as a scoring device. And a control unit which automatically administered 50-sec. trials separated by 10-sec. rest pauses was added to the experimental equipment.

A total of 36 Ss was used in the experiment. They were divided into two groups which differed in the length of the practice trial and in the length of the inter-trial rest period, but received the same total amount of practice and the same total amount of rest. The first of these two groups practiced for five min. and rested between practice sessions for six min. Each of the

five-min. practice sessions consisted of five 50-sec. trials separated by 10-sec. pauses during which the timer was read and reset. Since the *S* began tracing the target during the last few sec. of the inter-trial interval, a short rest pause of about five or six sec. between trials was allowed in addition to the six min. pause between sessions. The second group of *S*s practiced for 50 sec. and rested for 60 sec. between trials.

The *S*s were run in the experiment two at a time in alternation; i.e., while one *S* practiced, his partner rested and vice versa. After each 50-sec. trial, the practicing *S*'s score was read to him. Although no allusion was ever made to the fact, this arrangement meant that the *S*s were working in a man-to-man competition situation. It also meant that, for purposes of later analysis, the *S*s in the five-min. practice, six-min. rest group could be subdivided into two further groups: those who succeeded and those who failed in this competition. The inference that then seemed possible was that *S*s who failed in the competition would be more highly motivated to do well than those who were successful.

All *S*s were given 25 practice trials. The data obtained were scores on each trial in terms of the number of sec. the *S* kept the stylus in contact with the target.

Results and Discussion

Learning curves for the two experimental groups are presented in Fig. 1. Here the mean score for each group is plotted as a function of the ordinal number of the learning trial. The plot with the open circles is for the group which practiced and rested for 50 sec. and 60 sec. respectively. This curve (to which we shall refer as a spaced practice curve) is fairly typical for the acquisition of skill under favorable conditions of distribution. The curve which is plotted in five segments is that for the group which practiced

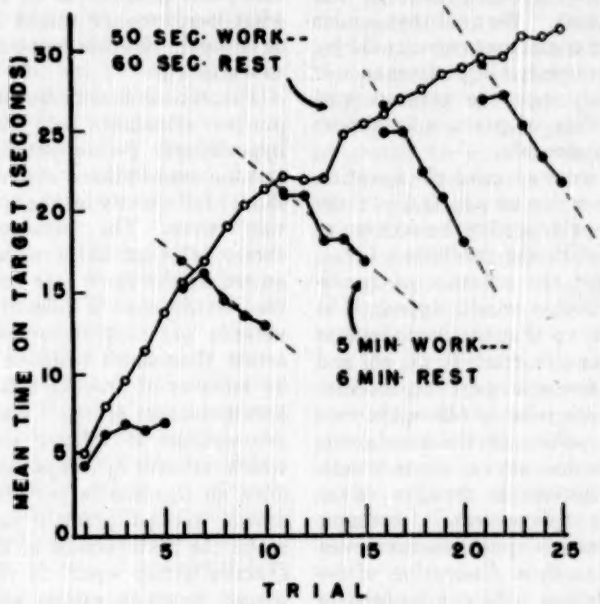


Fig. 1. Learning curves for the two experimental groups. The upper curve is for the group which practiced for 50 sec. with rest pauses of 60 sec. The curve plotted in five segments is for the group which practiced for five min. and rested for six min. Breaks in this function come at the points at which the rests were introduced. The straight lines are lines of best fit to the relatively decremental portion of the short curves. The solid black points are estimates of what the first post-rest performance score might have been had there been no need to warm up.

for five min. and rested for six min. The breaks in the function come at the points at which the six-min. rest pauses were introduced.

A first minor point to be made with respect to these curves is that they ought to cast some doubt on the suggestion that has been made that an appropriate measure in terms of which experiments on the distribution of practice might be analyzed would be a work-rest ratio. For the two groups whose performances are represented in this graph, this ratio would be the same ($\frac{1}{2}$). But from the dissimilarity of the curves, it is apparent that, if a work-rest ratio were to be considered in the analysis, nothing would be gained in simplicity.

A second point to be made is with regard to the general shape of the short functions. We note that, under the practice conditions represented by them, improvement in performance is made mainly between sessions and that the effect of practice is almost wholly detrimental.

Finally, we may raise the question of whether or not we can find in these curves any evidence for the accumulation of conditioned inhibition ($_sI_R$). Evidence for the presence of conditioned inhibition would appear if it could be shown that the initial points after rest (i.e., on trials 6, 11, 16, and 21) for the five-min. practice, six-min. rest group are reliably below the corresponding points on the continuous spaced practice curve, since conditioned inhibition is thought of as effecting a permanent decrement which depresses performance even after the complete dissipation of reactive inhibition. As can be seen in the graph, these first points are actually slightly below the spaced practice curve. They are not, however, significantly so. Student *t*-values for the significance of these four differences are .14, .47, .58 and 1.36, none of which approaches the 10 percent level of confidence. Furthermore, these first points after rest are depressed, as has been shown by Ammons (1), by a need to warm up. We have used Ammons' technique for estimating the amount of warm-up decrement, and have indicated in solid black circles estimated points which represent what performance on each of the first post-rest trials might have been had there been no need to warm up. The technique involves fitting the straight lines shown in Fig. 1 to the relatively decremental portions of the post-rest performance curves and extrapolating these lines backward to the first post-rest trial. The level of the line at the first post-rest point is taken as an estimate of what performance might have been at this point if there had been no need to warm up.

Theoretical considerations permit a precise statement of where these hypothetical points should fall. If certain assumptions are valid, they should fall exactly on the spaced practice curve. The assumptions are these: (1) that habit or skill is acquired at the same rate regardless of the conditions of distribution involved; (2) that it is performance rather than habit which is depressed by massing of practice trials and the accumulation of I_R. Under these assumptions it follows that a rest which removes I_R completely will produce an increase in performance following massed practice sufficient to make the performance of the massed practice group equal to that of the spaced practice group unless some conditioned inhibition has developed and produced a long-term decrement in performance. Reference again to Fig. 1 will show that all of the estimated first points after rest are very

near the spaced practice curve. This suggests that the massed practice group learned as well as the spaced practice group and that no conditioned inhibition had developed.

Now let us turn to the final point of this paper, an analysis of the data for the five-min. practice, six-min. rest group when they are broken down in terms of *Ss* who were successful and *Ss* who were unsuccessful in the man-to-man competition situation inherent in the experimental design. Here we will take as our starting hypothesis that *Ss* whose scores were lower than those of their partners are more highly motivated to do well in the learning task than *Ss* whose scores were relatively high. In addition, we will make the assumption that this difference in motivation is established in the first five-min. practice session and is maintained (though perhaps in a decreased degree) throughout practice. Thus, on the basis of mean scores on the first five practice trials, it has been possible to identify two subgroups within the five-min. practice, six-min. rest condition; (a) *Ss* whose scores were, on the average, lower than those of their partners (low group), and (b) *Ss* whose average scores were higher than those of their partners (high group).[1] Our prediction now is that *Ss* in the low group will tolerate more reactive inhibition than those in the high group because they are more highly motivated, and that, therefore, the *Ss* in the low group will consistently show a greater amount of reminiscence after the six-min. rests introduced at four places in the learning series. A graphical test

[1] A point of some importance is that, by this procedure, the two groups show considerable overlap in learning scores. Since the high or low designation is based on comparison with the *partner's* score, some of the *Ss* in the low group had considerably higher scores than some *Ss* in the high group.

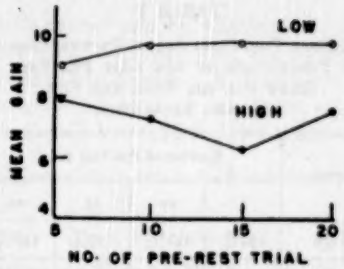


FIG. 2. Mean gain plotted as a function of the trial after which the gain occurred. The upper curve is for *Ss* whose mean scores on the first five practice trials were lower than those of their partners. The lower curve is for the *Ss* whose scores were higher than those of their partners on trials 1–5.

of this prediction is presented in Fig. 2 where mean gains for the high group and low group separately are plotted as a function of the trials after which these gains occurred. The differences are clearly in the expected direction. For purposes of statistical evaluation of the over-all difference presented in this graph, a mean gain score was computed for each *S* which was the arithmetic average of the four individual gains. A *t*-test of the significance of the difference between the mean of these values for the high and low group yields a *t*-value of 2.875 which is significant at the two percent level of confidence.

A similar effect can be shown when the post-rest score is represented as a percentage of the pre-rest score, thus giving an index of the relative increase in performance for the low and high groups. These percentages are presented in Table I. The percentage increase is always in favor of the low group.

In the preceding few paragraphs, we have been able to show that the low group is working under a greater amount of reactive inhibition than the high group. Our hypothesis has been that this difference is a result of

TABLE I

THE FIRST POST-REST SCORE REPRESENTED AS A PERCENTAGE OF THE LAST PRE-REST SCORE FOR THE HIGH AND LOW GROUPS SEPARATELY

Group	Number of Pre-Test Rrial			
	5	10	15	20
High	194%	148%	139%	148%
Low	334%	226%	182%	160%

an inferred difference in motivation. There is, however, an alternative hypothesis which needs to be ruled out before we can take the interpretation in terms of motivation very seriously. This alternative is that the greater amounts of reminiscence between sessions are simply a characteristic of poor learners in general and have nothing to do with motivation or with inhibition. Two lines of evidence appear to make this second hypothesis less plausible than it seems at first to be. The first of these is that, toward the middle and at the end of the experimental period, our group which was designated as low on the basis of performance on the first five trials is no longer low. On the very first learning trial, the high group's mean score *exceeded* that of the low group by 154 percent. On trials 6, 11, 16 and 21 the comparable percentages are 17, 16, 2 and 1. In other words, the relatively poor performance of the low group did not persist much beyond the first half of the learning session. At the same time, that there is a real difference in the amounts of I_R tolerated by these groups is indicated by the data presented in Fig. 2 and in Table I.

A second line of evidence which tends to rule out the suggestion that the greater inter-session increase is a function of poor learning *per se* is the following. If greater increases with rest were related directly to a low level of performance, there ought to be a negative correlation of some magnitude (for the total group of *S*s in the massed practice condition) between level of learning and amount of increase with rest. To test for this possibility, rank order correlations were run between mean performance scores within each session and the amount of gain between this session and the next. The rank-order coefficients for the four segments of practice where this computation was possible were as follows: mean score, trials 1–5 vs. gain on trial 6: + .26; mean score, trials 6–10 vs. gain on trial 11: + .27; mean score, trials 11–15 vs. gain on trial 16: −.24; mean score, trials 16–20 vs. gain on trial 21: −.10. Thus, there is only a tendency for the predicted negative correlation to occur, and that appears late in learning when, as we have seen, the high and low performers are no longer identical with our original high and low groups.

SUMMARY AND CONCLUSIONS

The results just presented have served to make two general points: (1) that, under the conditions of massed practice used in this experiment, no conditioned inhibition developed, and (2) that differences in motivation are probably a factor in determining the amount of reminiscence, presumably because an increase in motivation leads the *S* to tolerate a greater amount of reactive inhibition.

With regard to the first of these two points, probably no further comment is necessary. The results confirm those obtained in an earlier investigation designed to test this prediction more directly (3). The main value of the present findings in this connection stems from the fact that, here, a new

learning task and new practice conditions have been employed. Since these results are in essential agreement with the earlier ones, the suggestion is that our previous statement of the conditions under which ${}_{S}I_{R}$ develops (several minutes of uninterrupted practice) and the theoretical notion this suggests (a threshold of I_R which must be reached before ${}_{S}I_{R}$ begins to develop) are not limited to a particular learning situation.

In connection with the second hypothesis (that the amount of I_R an S will tolerate and therefore the amount of reminiscence possible depends partly upon the S's motivation) we must be more cautious. We have been able to show that the experimental findings are not incompatible with this hypothesis. And we have been able to show that one alternative explanation can be eliminated. But this does not mean that the hypothesis has been *proved*. As is usual in the progress of science, we have simply made this hypothesis more tenable.

(Manuscript received March 29, 1949)

References

1. Ammons, R. B. Acquisition of motor skill: II. Rotary pursuit performance with continuous practice before and after a single rest. *J. exp. Psychol.*, 1947, 37, 393–411.
2. Kimble, G. A. An experimental test of a two-factor theory of inhibition. *J. exp. Psychol.*, 1949, 39, 15–23.
3. Kimble, G. A. Performance and reminiscence in motor learning as a function of the degree of distribution of practice. *J. exp. Psychol.*, 1949, 39, 500–510.
4. Kimble, G. A., & Horenstein, B. R. Reminiscence in motor learning as a function of the length of interpolated rest. *J. exp. Psychol.*, 1948, 38, 239–244.

THE ROLE OF RECENCY IN LEARNING

BY R. H. WATERS AND JOHN G. REITZ

The College of Wooster

Introduction

The validity of recency as a principle of learning has been denied on both theoretical and experimental grounds. It might seem that further discussion of the preinciple is unnecessary. Two reasons may be given in justification of continued study of the principle: first, unanimity does not exist with respect to its complete lack of validity; the second is the distinction which has been made in theoretical discussions between learning and performance. This distinction suggests the possibility that recency may operate in performance, if not in learning, and it would therefore be desirable to test the principle in that area. Before starting that task, however, it might be well to review and evaluate some recent evidence favoring recency as a principle of learning.

In a recent paper Voeks has submitted the principle to an experimental test in an effort to discover its usefulness as a basis for predicting performance in a learning situation. Her results lead to the conclusion that such predictive value is high. "It was found that prediction of responses based upon postremity (the term used to replace recency and meaning the condition of being the last response made to a stimulus pattern) corresponds remarkably closely to future observed responses—not only for all groups but for all individuals, the predictions being correct for no individual less than 73 percent of the time and for most individuals more than 85 percent of the time. For each individual, the accuracy was significant considerably beyond the 0.1 percent level of confidence" (3, p. 608).[1]

These predictions were made in a maze learning situation in which the human subjects were given instructions to learn and in which their learning was apparently carried to a criterion of three perfect successive runs. As Peterson (1) has pointed out recency, or postremity, will lead to 100 percent correct predictions of the next response when learning is complete. This being obviously true, it is clear that, in a learning situation, the role of recency as a causal factor is difficult to establish beyond reasonable doubt to the contrary. Miss Voeks does not claim for postremity the status of a causal principle. Her formulation is as follows: "In a recurring situation, such as repeated trials in a puzzle-box or maze, the most probable response on any trial is the one last made to the stimuli present on that trial. If some transitory cause alters the response, this new response on any trial is the one predicted for the next trial on which that stimulus pattern occurs." (p. 495). In a later place she remarks, "Actually, by the present hypothesis, postremity itself is neutral, having no differential effect on the correct and incorrect responses" (p. 497). Postremity is, in other words, simply an index, a predictive device for anticipating what the next-succeeding response on the part of the subject will be.

Because her positive findings are supported by data obtained from re-

[1] In Voeks' paper a summary of much of the relevant literature is presented and hence such a review will be omitted here.

sponses made by subjects who had been instructed to learn the maze, the predictive validity of postremity is possibly limited to such situations. It also appears that postremity is a descriptive principle only, that the causes for the sequence of responses are other than postremity. It may be that such predictive value is of sufficient importance to give postremity some sort of scientific status, but it surely is insufficient to satisfy one who is looking for the factors and conditions that cause learning.

Statement of the Problem

Before these suggestions as to the significance and limitation of postremity can be considered more than hypotheses, a more detailed study of the precise relationship between postremity and subsequent responses in learning situations should be made. It was for this purpose that the study to be reported was undertaken. The question was: What is the predictive value of postremity or recency in a maze learning situation and what is the significance of such predictive value as it may be found to have? The experimental design was arranged to permit the test of another question: If the experimental conditions are altered on successive trials, how is the predictive value of recency affected?

Apparatus and Procdeure

The apparatus employed was a 15-choice-point, raised-wire, finger, revolving maze developed by Sharp (2). The maze pattern was mounted on sheet metal attached to the circumference of a bicycle wheel driven by an electric motor at an approximate speed of one revolution in 45 sec. A small opening permitted the subject to place his finger on the pathway and to see a small part of the pattern lying between choice points. At no time could the subject see the next choice point; he made his choice before it appeared in the opening. An incorrect choice was indicated by a small loop over the wire paths. The absence of such a loop indicated the correct path. The maze continued to revolve during the entire period of learning, thus preventing the subject from correcting any incorrect choice. Although the maze as constructed may permit as many as five alternate pathways at each choice point, in the present experiment only two paths were presented. Because of the structure of the apparatus, it was not necessary to blindfod the subject. He thus had both visual and tactual information of the correctness or incorrectness of his choice point reactions.

The results of the experiment are based on the records of 20 college students, naive with respect to maze learning, who were given 15 trials on the maze under conditions of massed practice. A sample of a raised wire finger maze was shown the subject before he began work, the starting point on the revolving maze was shown him, the manner in which the maze would revolve was described, and he was told he would be required to continue until he could go through the maze without error. Actually he was given but 15 trials, too few to allow most subjects to complete the learning.

A record was made of each choice made by the subject at each of the 15 choice points for each of the 15 trials. Following the procedure employed by Miss Voeks, the choice made on trial one, for example, at each choice point, was predicted for trial two, that made on trial two was predicted for trial three and so on. Thus, predictions were made for all trials save the first. By this method there were 4200 ($20 \times 15 \times 14$) individual predictions available for analysis.

In the conduct of the experiment the 20 subjects were divided into two groups of 10 each. The members of the first group were seated before the maze and followed the pathway with the index finger of the right hand for all 15 trials. The members of the second group worked under the same conditions for the first 10 trials and then were asked to change their position for each of the remaining five trials. On the 11th trial they were required to stand and use the left hand, to stand and use the right hand on the 12th trial, to sit again and use the left hand on the 13th trial, and to sit and use the right hand on the 14th and 15th trials. The introduction of these changes in position permitted a test of the effect of a shift in stimulating context on the predictive value of recency. This change, on Miss Voeks' hypothesis, should have reduced markedly the number of successful predictions.

Results and Discussion

In the presentation of the results those showing the number of success-

TABLE I

CORRECT PREDICTIONS BY RECENCY FOR GROUPS

Group	N	Mean % Correct	Range	Median % Correct	S. D. (N = 20)	Sigma Mean (N = 20)
I	10	70	55–87	70.5		
II	10	69	57–83	69.5	9.24	2.12

ful predictions for the entire performance will be exhibited first and later the analysis of the data to show the effect of change in stimulating conditions on such predictions will be presented.

The first question is this: Considering the data as a whole, what was the mean number of successful predictions? The data are contained in Table I. This Table shows the mean and median percent correct predictions, and the range for each of the two sub-groups, together with the sigmas of the distribution and of the mean for the group as a whole.

As the table shows, prediction based on recency of response is correct in about 70 percent of the cases. The medians are approximately the same, 70 percent, and the ranges of the two groups agree farily closely, from 55 to 87 percent of the predictions were correct for the individuals in Group I and from 57 to 83 percent were correct for the individuals of Group II. These values are sufficiently close to justify computing the sigma values in terms of the two groups combined. These sigmas are, standard deviation = 9.24, sigma mean = 2.12.

These values are not quite as high as those reported by Miss Voeks. She reports, for different groups, percentages of 82, 83 and 89 correct predictions. As later analysis will show, the major cause of this difference is probably to be found in the fact that her subjects were carried to mastery of the maze instead of through a fixed number of trials below the level of mastery. Other differences in the instructions given the subjects are also partly responsible. Miss Voeks' subjects were apparently allowed to correct incorrect choices in part of her study and all were asked to count the alleys and to announce the number of the choice point before making the choice. No correction of wrong choices was allowed in the present sutdy and no formal instructions to count or number the choice points were given. Miss Voeks' instructions would undoubtedly serve to produce more rapid learning and to direct her subjects' attention to cues that would facilitate such mastery. Her particular technique thus introduced factors other than mere recency of response. Nevertheless, the two studies are in essential agreement on the predictive power of recency in a learning situation.

We turn now to a more detailed analysis of the relationship between

TABLE II

PERCENT PREDICTIONS CORRECT PER TRIAL FOR GROUPS

Group	Trials 2	3	4	5	6	7	8	9	10	11	12	13	14	15
I	53	65	71	63	66	74	73	64	70	71	76	75	84	77
II	53	63	69	67	71	65	67	71	78	72	71	77	79	83

TABLE III

TOTAL AND PERCENT PREDICTED CORRECT CHOICES PER TRIAL FOR GROUPS

Group		2	3	4	5	6	7	8	9	10	11	12	13	14	15	Totals
I	Total	83	95	92	95	107	112	106	104	102	110	116	116	120	123	1154
	%	61	64	67	70	78	81	83	81	76	81	83	87	84	90	78
II	Total	81	91	91	91	100	100	104	108	106	108	113	104	120	119	1090
	%	56	64	69	67	73	73	76	78	78	77	80	79	82	89	75

the predictions correctly made and the learning scores. In Table II are shown the percent corrections for each of the second to 15th trials for each group.

It is clear from the data of this table that successful predictions based on recency increase over the series of learning trials. The increase is not without inversions but the trend is clear. The data suggest that correctness of prediction increases as learning takes place. That this is the case is brought out more clearly by the data of Table III.

This table shows that the subjects were learning the maze in that the number of correct choices increased with an increase in the number of trials. The increase in the number of correct choices is not uniform and learning is not complete for the group as a whole since in no case is the number of correct choices for any trial 150 as it would have to be to show that all members of the group went through that trial without error. The second row of figures for each group gives the percent of the correct choices predicted by recency for each trial. These values show clearly that successful prediction increases as learning goes forward. Had the subjects been held to some high criterion of mastery, recency predictions would have reached an accuracy of 100 percent. As has been said earlier, the higher predictive values found by Miss Voeks' are the result of the fact that her subjects were required to reach a criterion of three perfect successive runs.

This finding is not very startling; it is precisely that which was expected before conducting the experiment. It is the type of result that earlier theoretical analyses of the recency principle would predict. It does, however, serve to raise the question of the status of the principle of recency in relation to learning and to provide the ground for an evaluation of its predictive validity. It would seem clear that recency is not a causal factor in learning as is reinforcement for example. Its role is similar to that assigned by modern writers to frequency, exercise, or repetition, it gives the conditions causing learning a chance to operate. Upon what does the predictive value rest? Two answers suggest themselves. In the first place, knowing what response has been the more recent we can predict, on the basis of the assumption that factors causing learning were present at that time, that it will recur. Such a prediction is made on somewhat insecure grounds, since the organism may have learned to avoid that response in the future. Therefore, for prediction of the next response in terms of recency we need to know *what the organism is trying to do*. It is because the human subjects were trying to learn the maze that recency of response could be used to predict succeeding responses.

This interpretation is supported by

the second problem investigated in this experiment, the effect on the accuracy of prediction when the individual is asked to change his position from trial to trial. This was the condition under which the subjects of Group II worked during trials 11 to 15. These changes involved, as mentioned earlier, a shift from a sitting to a standing position and the reverse and the use of different hands on successive trials. Such changes in position involve changes in the stimulating conditions present on successive trials and, according to Miss Voeks' analysis, such changes should have been accompanied by a decrease in the accuracy of predictions based on recency. The relevant data are contained in Tables II and III. A review of these tables reveals little change in the percent of correct predictions for trials 11 to 15 from that found during trials 2 to 10. There is no more marked fluctuation during these trials than during the earlier trials when stimulating conditions were relatively more constant. The data do not support the notion that such changes in position and stimulation disturb the operation of the principle of recency nor the predictions based upon it. In terms of the data of Table III, particularly, we can say that the subjects were set to learn, were learning, and continued to learn in spite of shifts in bodily orientation. This fact supports the interpretation that successful prediction in terms of recency is a function of what the organism is trying to do in the situation.

There is another possible place for the principle of recency in behavior theory and that is in connection with performance. It may be that the principle has greater validity and meaningfulness when used as a predictive index for behavior in a non-learning situation. This has not been tested in the present experiment. It is questionable whether it will be found to have predictive validity in many such situations. Consider the act of walking. Certainly when one has moved his left leg forward it can safely be asserted that the next leg he moves will not be that one. This illustration raises the question of the units of behavior that are being predicted but it also implies that in order to make predictions of behavior one needs to know what the organism is doing.

Summary

To study the accuracy and significance of the predictive validity of recency in maze learning, 20 subjects were given 15 trials on a revolving, raised-wire, finger maze. Choices made at each choice point were predicted for the next choice at the same point. A total of 4200 individual predictions were analyzed. For the entire group approximately 70 percent of the predictions were correct. An increase in the percentage of correct predictions occurred with increased mastery of the maze. Requiring the subjects to change their bodily orientation during successive trials did not cause a decrease in the accuracy of predictions based on recency.

The results are interpreted to mean that recency is not a causal law of learning. It is, rather, a condition whose presence allows those factors causing learning to operate. It is also argued that successful prediction in terms of recency requires a knowledge of what the organism is trying to do. It is suggested that, without such knowledge, recency will not yield valid predictions in non-learning situations.

(Manuscript received April 22, 1949)

References

1. Peterson, J. Learning when frequency and recency factors are negative. *J. exp. Psychol.*, 1922, 5, 270–300.
2. Sharp, W. L. The relationship between speed and efficiency of learning on the constant speed finger maze. *J. exp. Psychol.*, 1939, 24, 86–94.
3. Voeks, Virginia W. Postremity, recency, and frequency as bases for prediction in the maze situation. *J. exp. Psychol.*, 1948, 38, 495–510.

THE RELATION OF STRUCTURAL CHANGES OF THE EYE AND VITAMIN A TO ELEVATION OF THE LIGHT THRESHOLD IN LATER LIFE

BY JAMES E. BIRREN, MALCOLM W. BICK, AND MARVIN YIENGST

The Section on Cardiovascular Diseases and Gerontology, National Institutes of Health, Bethesda, Maryland; and the Wilmer Ophthalmological Institute of Johns Hopkins Hospital and University, Baltimore, Maryland •

A decrease in visual sensitivity as measured by the light threshold has been shown to occur in later life (2). Although the tendency to reduced pupil size in the elderly affects the amount of light reaching the retina, a significant age change in threshold has been demonstrated irrespective of pupil size (2). It is reasonable to suspect that the decrease in visual sensitivity with age may be related to structural changes in the eye (4). However, there is a lack of correlated studies in which functional tests have been applied to eyes examined for structural changes.

The present study was undertaken to determine, if possible, some of the factors associated with the age decline in the sensitivity of the eye. For this purpose measurements of the light threshold were related to the findings of a complete ophthalmological examination of each subject. Such examination included retinoscopic and static refraction, complete external and ophthalmoscopic examination and in addition in some cases biomicroscopic examination.

Since there also existed the possibility that vitamin A may have an important limiting effect on the visual function of the older person, a study was made of the relation between the light threshold and serum levels of vitamin A and the response of the light threshold to vitamin A supplementation.

• The assistance of Miss Charlotte Fox and the cooperation of the William Hodson Recreation Center for the elderly, New York, is gratefully acknowledged.

Procedure

Subjects.—A total of 109 male subjects aged 40–83 was used in the present study: 56 from the Baltimore City Hospitals home for the indigent aged, 51 members of the William Hodson Recreation Center for the elderly in New York, and 2 laboratory technicians.

The Baltimore subjects were all ambulatory and did not require hospitalization. Regulations for admission require that individuals be able to care for themselves physically but lack financial means to maintain themselves in the community. The New York group was still resident in the community and attended a voluntary recreation center for the elderly. The principal financial support of this group was state and federal old age assistance.

Light threshold.—The procedure used for measuring the minimum light threshold has been described previously (2, 5) and is only summarized here.

A Hecht Shlaer adaptometer was used in all threshold measurements. The test field was circular 3° and lay 7½° nasal to the fixation point, i.e., the temporal retina was stimulated. It was exposed with violet light (below 460 mμ, Corning 511) for 0.2 sec. Light thresholds were measured after three min. of light adaptation at a brightness of 479 ml. Measurements were made at one-min. intervals for the first 10 min. after preexposure to bright light and at two-min. intervals thereafter until 30 min. had elapsed. The mean of the last two determinations was used as the subject's threshold.

Pupils were dilated with five percent euphthalmine administered two or three times within the hour prior to testing in order to eliminate the miosis in older subjects. Light threshold values were converted to a standard five-mm. pupil (2).

Vitamin A analysis.—Analyses for serum vitamin A were made on 30 subjects aged 30–80 years, each of whom had been measured for light sensitivity. Fasting blood was used and the analysis for vitamin A was made with the tech-

nique of Bessey *et al.* (1) modified as a macromethod with venous samples (10). Correlations were determined for the interrelations of vitamin A, light threshold, and age.

In addition to the above study, 12 of the subjects were fed 100,000 units of vitamin A daily by mouth for a period of 38 days. Four of these subjects were continued on the same dosage of vitamin A for an additional 38 days, i.e., for a total of 76 days. Seven subjects were given a placebo for 38 days paralleling the vitamin A group. Determinations of light threshold were made before the supplementary vitamin A and within a week after the last dose.

Visual acuity.—The visual acuity of each subject was determined with undilated pupils using a pin hole. In addition, each subject while dilated was refracted to determine the proper lens correction. The best visual acuity was used in the analysis.

Visual acuity was measured at a distance of 20 feet with a Snellen E chart illuminated by 52 foot candles.

Criteria of structural changes.—Each subject was examined with the ophthalmoscope while his pupils were dilated. The examination was carried out in a predetermined order and the absence or degree of various conditions noted on a check list. It is difficult to estimate the exact degree of a structural change so that it may be treated in a quantitative manner. The criteria are arbitrary and essentially rest upon comparison with the picture of the average young adult eye. A five-point scale was originally adopted in the present study for rating ocular changes, zero representing no change and four the maximum change 'ever observed.' So few cases were classified in the higher category that only three intervals are reported in the results, i.e., '0' no evidence, '1' some evidence, and '2' clearly evident. These units are arbitrary and do not necessarily represent equal steps. The following descriptions represent those factors entering into the rating of ocular changes in the present study.

1. Macular changes.—The macula in the young adult is darker than the surrounding retina and is marked by a bright-light reflex in the center. This bright-light reflex appears to be an accentuation of the general retinal sheen noted in young persons. The retinal sheen diminishes in middle age while the foveal light reflex normally persists longer than the surrounding retinal glossiness. The disappearance of this foveal reflex may be followed by a fine granular appearance of the macular area which appears to become more coarse in older eyes. Also observed are small reddish pigment-free spots which became larger and coalesce, and may eventually give way to the frank 'beaten copper' appearance of macular dystrophy. Subsequent to this, there may appear an irregular clumping of pigment with white or yellow spots in the macula. The advanced senile macular lesions have been well described and classified by Lloyd (7).

2. Retinal changes.—The early retinal changes resemble those seen in the macular area. Subsequent to the diminution of the retinal sheen, the retina develops a granular quality. Such granularity may give way to a 'moth-eaten' appearance followed by a clumping of pigment and the gradual exposure of the choroidal vascular pattern. These changes may be associated with the occurrence of white and yellow spots of various sizes, scarring, exudation, and hemorrhages. In the present study, the characteristic picture of choroidal sclerosis was considered to be an advanced retinal change.

3. Lenticular changes.—Subjects with advanced lenticular changes were eliminated because observation of the fundus could not be

TABLE I

FACTORS DETERMINING RATING OF STRUCTURAL CHANGES

Macular Degeneration

Degree	Foveal Reflex	Changes in Pigmentation	Gross Macular Lesions
0	Present	None	None
1	Absent	Slight	None
2	Absent	Moderate	None
3	Absent	Marked	None
4	Absent	Marked	Present

Retinal Degeneration

Degree	Granular Appearance	Changes in Pigmentation	Gross Retinal Lesions*
0	None	None	None
1	Present	None	None
2	Present	Present	None
3	Present	Moderate	Present
4	Present	Marked	Present

Lenticular Sclerosis

Degree	Blurring of Retinal Details	Manifest Lens Opacities
0	None	None
1	Slight	Slight
2	Moderate	Definite
3	Marked	Marked

* This includes a variety of conditions such as gross scarring, atrophy, colloid degeneration, exudation, hemorrhages, etc.

accomplished nor could an estimation of retinal function be made accurately. Changes in the crystalline lens of the eye were observed with the +10.0 diopter lens of the electric ophthalmoscope. Lens opacities appear dark against the red background of the fundus. By focusing on the retina the clarity of fundus details gives an indication of the transmission of light through the lens. As the sclerosis of the lens increases, the retinal image becomes more yellow and less distinct as a result of scattering of light and the greater absorption of the shorter wave lengths. The rating of the amount of sclerosis was based upon the color and distinctness of the fundus picture as well as the presence of manifest opacities.

Outline of structural evaluations.—Table I contains the general factors determining the rating of structural changes.

Results

Structural factors.—Degenerative changes in the eyes of subjects beyond the age of sixty years were found in more than half of the subjects (Table II). Of all eyes beyond the age of sixty, only 30 percent were free of retinal degeneration, 19 percent were free of macular degeneration, and 23 percent were free of lenticular sclerosis.

Such changes appear to be distributed symmetrically; i.e., they appear in both eyes in approximately the same degree at the same age. The correlation between the right and left eyes for retinal degeneration, macular degeneration and lenticular sclerosis were 0.73, $N = 88$, 0.69, $N = 102$, and 0.63, $N = 103$. These correlations are contingency coefficients based on 3×3 tables and are uncorrected for grouping.

Significant differences in the light threshold were found between the older subjects with and without retinal degeneration (Table III). The mean light threshold for subjects *free* from retinal degeneration was 2.90 log $\mu\mu$l. for the 60–69 year group and 3.16 log $\mu\mu$l. for the 70–83 year group. The means for the subjects *with* retinal degeneration were 3.25 and 4.11 log $\mu\mu$l. for the corresponding groups. These differences are significant at the three and one percent levels respectively.

Within the 70–83 year group the

TABLE II

Frequency of Ocular Defects by Age Groups*

	40–49		50–59		60–69		70–83	
	N	%	*N*	%	*N*	%	*N*	%
Retinal Degeneration								
Without	22	100	20	83	18	36	21	27
With	0	—	4	17	32	64	57	73
Total	22	100	24	100	50	100	78	100
Macular Degeneration								
Without	19	86	13	46	25	37	5	6
With	3	14	15	54	42	63	84	94
Total	22	100	28	100	67	100	89	100
Lenticular Sclerosis								
Without	18	82	18	64	22	33	14	16
With	4	18	10	36	44	67	75	84
Total	22	100	28	100	66	100	89	100

* Combined right and left eyes

TABLE III

MINIMUM LIGHT THRESHOLD (in log μμl.) AS A FUNCTION OF AGE AND RETINAL DEGENERATION

Retinal Degeneration		Age 40–49	50–59	60–69	70–83	Total
0	M	2.66	3.06	2.90	3.16	2.94
	σ_M	0.15	0.23	0.12	0.12	0.08
	N	11	10	10	10	41
1+	M	—	3.09	3.30	3.47	3.36
	σ_M	—	—	0.11	0.22	0.12
	N	0	2	12	13	27
2+	M	—	—	3.14	4.63	4.27
	σ_M	—	—	0.20	0.39	0.26
	N	0	0	5	16	21
Total	M	2.66	3.07	3.12	3.86	
	σ_M	0.15	0.20	0.08	0.18	
	N	11	12	27	39	

TABLE IV

MINIMUM LIGHT THRESHOLD (in log μμl.) AS A FUNCTION OF AGE AND LENTICULAR SCLEROSIS

Lenticular Sclerosis		Age 40–49	50–59	60–69	70–83	Total
0	M	2.66	3.20	3.14	4.55	3.24
	σ_M	0.16	0.25	0.16	0.70	0.18
	N	11	9	10	6	36
1+	M	—	2.68	2.97	3.59	3.33
	σ_M	—	—	0.11	0.20	0.14
	N	—	3	11	22	36
2+	M	—	—	3.88	3.95	3.92
	σ_M	—	—	0.39	0.22	0.20
	N	—	—	8	12	20
Total	M	2.66	3.07	3.28	3.84	
	σ_M	0.16	0.20	0.14	0.17	
	N	11	12	29	40	

mean change in light threshold associated with retinal degeneration was greater than that associated with age alone. However, since elevation of the light threshold is found in persons free from retinal degeneration, it might reasonably be expected that other factors may also influence the aging of the light threshold.

The presence of lenticular sclerosis, to the extent included in this study, was not associated with elevation of the light threshold (Table IV).

In the 70–83 year group, the mean of persons without lenticular sclerosis was actually higher than that for persons with sclerosis, although the difference was not statistically significant ($P > .10$). This suggests that there is little correlation between retinal degeneration and lenticular sclerosis. The coefficient of contingency for the relation between lenticular sclerosis and retinal degeneration in the 70–83 year group was low, i.e., 0.35. In these data the relation cannot be regarded as significantly greater than zero.

Macular degeneration was associated with elevated light thresholds in the 70–83 year group (Table V). In the 60–69 year groups no such relation appeared. This suggests that the coexistence of retinal and macular degeneration in the 70–83 year group

TABLE V

MINIMUM LIGHT THRESHOLD (in log μμl.) AS A FUNCTION OF AGE AND MACULAR DEGENERATION

Macular Degeneration		Age 40–49	50–59	60–69	70–83	Total
0	M	2.61	2.85	3.63	2.86	2.84
	σ_M	0.16	0.18	—	—	0.13
	N	10	6	3	2	21
1+	M	3.14	2.96	3.15	3.22	3.16
	σ_M	—	0.17	0.11	0.12	0.07
	N	1	4	17	16	38
2+	M	—	2.60	2.88	4.40	4.03
	σ_M	—	—	0.07	0.26	0.24
	N	—	2	4	20	26
Total	M	2.65	2.84	3.16	3.82	
	σ_M	0.16	0.11	0.10	0.18	
	N	11	12	24	38	

was responsible for the elevated thresholds. Upon inspection of the data it was found that the elevation of the light threshold associated with macular degeneration in the 70–83 year group was primarily contributed by 12 individuals who had both marked retinal and macular degeneration. The present data do not permit further analysis of this issue.

These results indicate that poor visual function, i.e., the minimum light threshold in the elderly, is associated with observable degenerative changes in the eye. Only one subject beyond the age of 70 was free of both retinal and macular degeneration (he had slight nuclear sclerosis). The light threshold of this subject was 2.7 log $\mu\mu$l. which is about the mean for teen age children or young adults (2).

The correlation (r) between the light threshold and visual acuity was 0.49, $N = 91$. This somewhat low correlation indicates that it is not possible to evaluate the visual efficiency of an individual simply from the measurement of visual acuity alone. Visual acuity declined with age in the present subjects; the correlation between age and acuity was −0.56 (contingency coefficient, $N = 212$ eyes). In the age group 70–83, only 20 eyes of a total of 117 eyes refracted had acuity of 20/30 or better.

Vitamin A.—The mean fasting serum vitamin A level was 48.7 μgm. percent, $\sigma = 14.8$. No relation appeared between the vitamin A levels and either age or light thresholds (Table VI). This is in agreement with the findings that: (a) there is no shift in plasma vitamin A with age (6), (b) in young non-deficient subjects there is little relation between plasma vitamin A and the light threshold (3), and (c) there is no age change in gastrointestinal absorption of vitamin A (10).

TABLE VI

CORRELATION (r) BETWEEN THE LIGHT THRESHOLD, FASTING VITAMIN A PLASMA LEVEL, AND AGE, FOR 30 SUBJECTS AGED 30 TO 80 YEARS

	Light Threshold	Age
Vitamin A	−0.15*	−0.20*
Light Threshold	—	0.61

* Not significantly different from zero

Since serum vitamin A levels do not necessarily reveal tissue levels or liver storage, a therapeutic regime was undertaken as outlined in the procedure. No significant change was produced in the light threshold with 38 days of 100,000 I.U. daily in 12 subjects or in 4 subjects given the same dose for 76 days (Table VII).

Serum vitamin A determinations were made before administration and 48 hours after the last dose of nine of the subjects receiving vitamin A for 38 days. The mean initial level was 56.1 μgm. percent and the mean level 48 hours after the last dose was

TABLE VII

CHANGES IN LIGHT THRESHOLD (in log $\mu\mu$l.) AFTER 38 DAYS AND 76 DAYS OF 100,000 I.U. OF VITAMIN A DAILY; AND AFTER 38 DAYS OF A DAILY PLACEBO

	Vitamin A 38 Days			Vitamin A 76 Days			Placebo 38 Days		
	Pre-	Post-	Dif.	Pre-	Post-	Dif.	Pre-	Post-	Dif.
Mean	3.07	2.89	−0.18	3.02	3.05	+0.03	2.97	3.02	+0.05
N, t, P	12; 1.32; .20			4; .40; .70			6; .31; .70		

71.1 μgm. percent. This increment is significant statistically ($P = .02$).

While it is possible that continued administration of daily supplementary vitamin A might have lowered the light threshold, it does not appear likely. Steven (8) advanced the criterion of a pre-existing deficiency by lowering the light threshold 0.3 log μμl. or more in two to three weeks of vitamin A supplementation. This was not observed in the present group even in the case of two subjects who had high thresholds, i.e., 3.4 and 3.3 log μμl., and who had been given 100,000 I.U. of vitamin A per day for a total of 76 days.

Such findings do not exclude the possibility of impaired utilization of vitamin A in the elderly or a relation between impaired utilization of vitamin A and structural changes which in turn affect the light threshold.

Discussion

The more prominent features of the present results are: (a) the prevalence of degenerative changes in the older eye, (b) the relation of the light threshold to retinal degeneration, (c) the lack of relation between serum vitamin A levels and the light threshold, and (d) the apparently complex nature of the aging of the eye. The last of these has perhaps the greater theoretical implications whereas the first is of more immediate practical significance. The differences in the rates of aging of the various structures of the eye implies that a study of aging of a visual function soon becomes one of partialling out the influence of numerous factors each of which is correlated with age. An example of a differential age effect is that of nuclear sclerosis. Presumably nuclear sclerosis is one of the earlier signs of aging of the eye yet its presence would affect visual acuity more by virtue of the scatter of light than it would the minimum light threshold. Another example of the diverse nature of aging of the eye is the correlation between nuclear sclerosis and retinal degeneration. Were it not for the low relation between these two processes no benefit would be derived from cataract removal since the retina would also show degenerative change. In the case of an older eye with some nuclear sclerosis, the correlation of visual acuity and the light threshold in the same individual would depend upon the additional factor of the presence of degenerative changes in the retina and macula.

The concept of the 'normal' old eye must necessarily refer to a small population if an eye free of degenerative changes is referred to. Similarly the use of a single test such as visual acuity to estimate the visual capacities of older subjects is of dubious validity.

Summary and Conclusions

1. Studies were made of the relation of the minimum light threshold to changes in the retina and transmission system of the eye. A total of 109 male subjects, aged 40–83, was observed. The light threshold was measured with the Hecht Shlaer adaptometer. The corrected visual acuity was determined by refraction for each subject. Observations of the appearance of the retina, macula, and transmission system were made with an ophthalmoscope.

2. In addition to a significant age decline in visual sensitivity as revealed by the light threshold, there was observed a high prevalence of changes in the retina, macula, and transmission system of the older eye. The light threshold of the older subjects was correlated with the presence of retinal

degeneration. None of the 39 subjects age 70–83 years was free of retinal and macular degeneration and lenticular sclerosis; only one subject was free of both retinal and macular degeneration. The light threshold of this subject was the lowest of the group and was at the mean level for young persons. The results suggest that it is difficult to generalize as to the efficiency of the older eye from one functional test such as acuity or light threshold, or to use any single observation to secure a 'normal' or disease-free older population in studies of aging.

3. Analyses for serum vitamin A were made on 30 subjects, aged 30–80 years, each of whom had been measured for light threshold. In addition, 12 of the subjects were given 100,000 I.U. of vitamin A daily for 38 days and 4 subjects were given the same dosage for 76 days; 7 subjects were given a placebo for 38 days. Determinations of the light threshold were made before and after the vitamin or placebo series. Results of vitamin A determinations indicated a lack of age change in serum vitamin A and a lack of correlation between the age change in light threshold and serum vitamin A. Daily supplementation with 100,000 I.U. of vitamin A produced no significant change in the light threshold of 12 subjects in 38 days or in 4 subjects in 76 days. Such findings should not be interpreted as indicating a lack of importance of vitamin A in the photochemical cycle. They suggest that the age changes in availability of vitamin A in the blood are minimal and uncorrelated with the light threshold.

(Manuscript received May 19, 1949)

References

1. Bessey, O. A., Lowry, O. H., Brock, M. J., & Lopez, J. A. The determination of vitamin A and carotene in small quantities of blood serum. *J. biol. Chem.*, 1946, 166, 177–188.
2. Birren, J. E., Bick, M. W., & Fox, Charlotte. Age changes in the light threshold of the dark adapted eye. *J. Geront.*, 1948, 3, 267–271.
3. Bodansky, O. L., & Haig, C. The comparative value of the blood plasma vitamin A concentration and the dark adaptation as a criterion of vitamin A deficiency. *Science*, 1941, 94, 370–371.
4. Friedenwald, J. S. The eye. In: E. V. Cowdry, ed. *Problems of aging*. Baltimore: Williams & Wilkins, 1942, pp. 535–555.
5. Hecht, S., & Shlaer, S. An adaptometer for measuring dark adaptation. *J. optic. soc. Amer.*, 1938, 28, 269–275.
6. Kirk, E., & Chieffi, M. Vitamin studies in middle-aged and old individuals: 1. The vitamin A, total carotene and α and β carotene concentration in plasma. *J. Nutrit.*, 1948, 36, 315–322.
7. Lloyd, R. I. Senile macular lesions. *Amer. J. Ophthal.*, 1948, 31, 1135–1146.
8. Steven, D., & Wald, G. Vitamin A deficiency; A field study in Newfoundland and Labrador. *J. Nutrit.*, 1941, 21, 461–476.
9. Wald, G. Visual systems and the vitamins A. In: H. Kluver, ed. *Visual mechanisms*. Lancaster: Cattell Press, 1942, pp. 43–71.
10. Yiengst, M. J., & Shock, N. W. Effect of oral administration of Vitamin A on plasma levels of Vitamin A and carotene in aged males. *J. Geront.*, 1949, 4. (In press)

THE DESTRUCTION OF THE MÜLLER-LYER ILLUSION IN REPEATED TRIALS: I. AN EXAMINATION OF TWO THEORIES

BY WOLFGANG KÖHLER AND JULIA FISHBACK

Swarthmore College

In their investigation of figural after-effects, Köhler and Wallach mention the fact that such effects may deserve the psychologist's attention not only as interesting perceptual phenomena but also for methodological reasons (8, pp. 348 ff.). In perception and in other parts of psychology, we generally repeat experiments with given *Ss*, because in this fashion we obtain reliable measures of scatter, and so forth. Quite frequently, the number of individual trials is considerable. The authors point out that under such circumstances a fair amount of satiation is likely to develop, and that, as a consequence, later results may sometimes no longer refer to precisely the same situation as was to be investigated.[1] This reasoning also applies to experiments in which several situations rather than one are given in repeated cycles. In this case, the satiation established in each particular situation could affect further results not only with this situation but also with all others. Whether experimentation refers to one or to several situations, if repetition of trials really had such effects, and if the nature of the effects were not recognized, misinterpretations of results would, of course, be unavoidable.

As a special example, K. and W. discuss the fact that some well-known illusions tend to disappear when, with given *Ss*, their size is measured in ever-repeated trials. It is still widely believed that such illusions represent errors of judgment rather than actual distortions of visual objects. From this point of view, it may seem natural to assume that, when *Ss* deal with the illusion patterns in a great many trials, they gradually learn to avoid those errors, and that the illusions disappear for this reason. This is also the explanation which is now commonly given, and which is clearly implied when experiments of this kind are called 'practice experiments' (11).

The most interesting experiments in this field refer to the Müller-Lyer illusion. They have been done by Judd (4, 5), by Seashore (10) and by Lewis (9). K. and W. discuss the same pattern. They explain that its two parts are subject to different degrees of satiation, and that, as a result, the size of the illusion must gradually decrease when the pattern is repeatedly inspected. Consequently, they say, the disappearance of the illusion may be caused by satiation rather than by some form of learning. They might have added that any explanation which attributes the destruction of the illusion to learning has to face a major difficulty. The illusion has sometimes been destroyed when *Ss* did not know that they were dealing with an illusion pattern, and when they

[1] By satiation the authors mean the changed condition in which a part of the neural medium is left when the process of a perceptual object has occurred in this part for some time (8, p. 276). Occasionally they use the same term for the alteration of the tissue as a *development* rather than as an established fact. It is assumed that figural after-effects are observed when the processes of test objects react to the satiated condition of the tissue.

were not told how their results compared with the objective situation. It is not immediately clear how, under these conditions, their judgments could be greatly improved by learning.

A more detailed investigation seemed to us advisable for several reasons. In the *first* place, if it were actually proved that the destruction of the illusion in repeated trials is caused by satiation, then the general warning given by K. and W. would be justified. As a matter of principle, any experiments in which certain situations are repeatedly given might be influenced by the accumulation of satiation—provided, of course, that the materials in question are at all capable of establishing this effect, and of thus altering their own character-acteristics. *Secondly*, we were interested in the nature of the Müller-Lyer phenomenon and similar illusions. It seemed to us probable that the distortions observed in such patterns are effects of interaction among the processes which underlie the various parts of the patterns. But, on the assumption that it is learning which makes the illusions disappear in repeated trials, their indirect interpretation as results of faulty observation would appear to be fairly plausible. If it were shown that repetition destroys the illusions by means of satiation rather than of learning, one could no longer argue in this fashion. Many visual patterns, the appearance of which has never been suspected of being influenced by mistaken judgments, are strongly affected by satiation.

While, originally, our experiments were meant to clarify only these issues, we soon obtained results which we had not expected to find, and which may actually be more important than the answers to our initial questions. These results refer to the way in which satiation develops in repeated trials, and to its persistence once it is well-established.

K. and W. do not give an explanation of the M.-L. illusion as it appears before the pattern is subjected to 'practice' experiments. In the present connection, no such explanation is needed if it is assumed that satiation affects this pattern according to the same rules as hold for satiation in the case of any other patterns. It must be shown, however, that in the illusion figure these rules are actually followed. In this respect, the authors argue as follows: Satiation is always particularly strong in areas which are limited by converging lines or contours. The surfaces limited by the angles of the illusion figure are such areas; hence they must be more intensely satiated than other areas, most particularly, than the regions toward which the corners of the angles point. The lines of each angle will, therefore, soon lie between two areas, one of which is far more strongly affected than the other. (Cf. Fig. 1, in which the regions of maximal satiation are marked S.) Under these circumstances, the angles must be displaced into the less satiated regions, i.e., in the direction of their apexes. It follows that in the right half of Fig. 1 the horizontal distance must grow, and that in the left half the opposite must happen. Thus satiation must gradually reduce the size of the illusion. Actually, K. and W. report that, after prolonged inspection of the pattern, the illusion is much less conspicuous, even if during the inspection period no attempts have been made to compare the crucial distances.[2] This observation is concerned with a change of the illusion as such; it gives no independent proof that, in this pattern, satiation really develops as it is supposed to do. A further observation is not subject to this criticism. Not only the interior of the angles must be satiated, but also, even if less, the middle parts of the regions between the angles. Here again, however, satiation cannot be symmetrical. The region on the right side

[2] 8, p. 349

of Fig. 1 is much more 'enclosed' by its angles than is the corresponding region on the left side. Consequently, central parts of the right region will be more strongly satiated than the more open region on the left side. K. and W. tested this conclusion by showing two equal objects in the central parts of both regions when the illusion pattern had been inspected for some time. They found that the test object placed in the right area was far more strongly affected than its partner.[3] A third observation which refers to the same question will be mentioned later in connection with similar experiments of our own.

The thesis that, in a long sequence of trials, satiation rather than learning destroys the illusion may seem to be open to an objection. In earlier investigations, *S*s were not generally asked to fixate any particular point when they compared the two crucial distances. Under these conditions, their eyes probably moved in both directions from the point at which the distances meet. K. and W. do not feel that their thesis becomes inapplicable for this reason. It is true that so far no experimental evidence is available as to the pattern of satiation which develops when, in retinal and cortical terms, the lines of a figure repeatedly move one way and the other. But in each position of the M.-L. pattern, satiation affects certain areas rather than only its lines. If, statistically speaking, the eyes move more or less symmetrically from the center of the figure toward its ends, then summation of the satiation patterns corresponding to their various positions is likely to establish maximal satiation in the places in which it would be maximal if the center were fixated. In this premise, the illusion could be destroyed by satiation even when *S*s are not asked to fixate the center. To clarify this matter, one might try to compute the pattern of satiation which would gradually result under such conditions. We have not tried to make this computation, because it would have to be based on idealizing assumptions which may not be acceptable. It seems preferable to examine the problem experimentally. A test which may answer our question will be mentioned later.

FIG. 1. A Müller-Lyer figure used in the destruction of the illusion

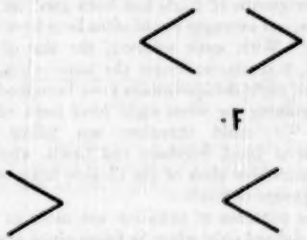


FIG. 2. A second Müller-Lyer figure used in the destruction of the illusion

The illusion patterns which we used were shown as black lines drawn with India ink on white cardboard. We worked with two versions of the M.-L. figure (Figs. 1 and 2), approximately the same as have been used by Lewis. But since results proved to be independent of the difference, we need not generally refer to it. The lines of the angles were 1 mm. wide and 2.7 cm. long. The size of the angles was 60°. No actual connecting lines were drawn between the corners of the angles. The illusion is at least as great without these lines as it is in their presence, and the judgments of our *S*s proved to be quite reliable when the distances were left empty. Throughout, the distance which appears shorter was kept constant at a length of 7 cm., and the other distance was varied by shifting the angle on its left side. Displacement of this angle was made possible by drawing its lines on the edge of a movable second screen which was pressed against the left side of the stationary screen. In the crucial region, the edge of the movable screen was cut so as to give the proper angle. We need not describe the simple device by means of which this screen (and the black angle on its edge) could slowly and evenly be shifted in one horizontal direction or the other. Since the device was hidden behind the screens, *E*'s operations remained invisible. Experiments were done in daylight; but lamps, shielded on the subjects' side, threw light upon the field in such a fashion that, apart from the black angle on its edge, the

[3] 8, p. 278

contour of the movable screen was barely perceptible when *S* fixated the center of the field. *S* observed from a distance of nine feet.

In measuring the size of the illusion, we used the method of limits. Trials in which the variable distance was first too long alternated with trials in which it was first too short. Apart from the usual deviations related to this difference of direction, results were remarkably constant at a given time. We soon decided that averages computed from four measurements, two in each direction, could be regarded as representative of the size of the illusion at a given point. In fact, if larger groups of trials had been used for this purpose, our averages would often have been misleading. With some subjects, the size of the illusion is surely no longer the same when, for instance, eight measurements have been made in the beginning, or when eight have been added later. We could therefore not follow the example of Judd, Seashore and Lewis, who derived successive sizes of the illusion from much larger groups of trials.

Since patterns of satiation are more or less clearly defined only when *Ss* fixate given marks, we prescribed fixation of the points (F) indicated in Figs. 1 and 2.

As to the schedule according to which trials were given during one session, and separate sessions on different days, we decided to use a flexible program. The temporal distribution of trials has so strong an influence on the decrease of the illusion that with certain rigid programs its destruction may actually be prevented. We varied our schedule as seemed most convenient for a study of such facts. In the following, discussion of this point will be postponed until simpler facts have been reported.

Other authors have found that the initial size of the illusion, i.e., its size before 'practice,' varies within a wide range from one *S* to another. In our experiments, this was also the case. We measured the initial illusion with 13 *Ss* altogether, and obtained as the smallest amount a distortion of 8 percent, as the highest, of 46 percent. The mean for the whole group was 27 percent; in other words, when subjective equality was established, 7 cm. in one part of the pattern appeared on the average equal to about 5 cm. in the other part. Although, with individual *Ss*, distortions of this size have been reported before, averages do not seem to have been so high in any previous investigation. We may have found greater distortions because our *Ss* were not allowed to move their eyes, because the M.-L. illusion tends to be greater when the corners of the angles are not connected by lines, and also because in all earlier investigations the initial size of the illusion has been derived from a great many individual measurements. For reasons which we have just mentioned, this procedure is likely to give values which are too low.

With five *Ss* (A–E) measurements were continued until the illusion had disappeared.[4] When this point was reached, the number of individual measurements previously made by the *Ss* were as follows:[5] A, 20; B, 24; C, 128; D, 144; E, 200. With the same *Ss*, the initial values of the illusion (in percent) had been, respectively: A, 8; B, 22; C, 15; D, 34; E, 46. If we except the relative positions of *Ss* B and C, the rank order given in the first series of figures is repeated in the second series. With only five *Ss*, one cannot, of course, prove that the number of trials needed for the destruction of the illusion grows with its initial size, although such a correlation may appear to be plausible. A perfect correlation cannot be expected under any circumstances. For instance, in a given *S* a high initial value of the illusion may be combined with exceptionally easy learning, if this is the factor which destroys the illusion, or with unusually strong satiation, if this is the responsible process. Quite

[4] At a distance of nine feet, comparisons such as those done in our experiments are not objectively reliable when differences are reduced to less than one mm. We therefore regarded the illusion as destroyed when the average of four measurements fell below this limit.

[5] These figures refer to numbers of individual measurements in a strict sense. At the time, from 5 to 50 averages had been computed from groups of four trials each.

apart from such considerations, the results of our *Ss* cannot be compared in a strictly quantitative sense, because the schedule of experimentation was not the same in all cases.

With our five *Ss* the illusion disappeared in far fewer trials than had been found necessary in earlier investigations. Judd, who worked with three *Ss*, had to give them 600, 800, and 980 trials to achieve the same end. Seashore and Lewis did not use the method of limits; but the total number of comparisons made by their *Ss* when the illusion disappeared was undoubtedly of the same order of magnitude. For our *Ss*, the average was only 103 trials. On the assumption that the illusion is destroyed by satiation, it may seem natural to explain the difference in this fashion: When *Ss* are allowed to move their eyes, a pattern of satiation which compensates for the illusion can be established only statistically, and therefore slowly. On the other hand, when fixation is prescribed, such a pattern will at once begin to develop in the areas in which it must be located if the illusion is to disappear. The illusion will therefore decrease more rapidly in the latter situation. Actually, this account of the growth of satiation, when *Ss* fixate, is probably not quite correct. Nevertheless, so far as *Ss* A and B are concerned, the explanation fits the facts. With these *Ss* the illusion was destroyed within less than an hour, i.e., in a single experimental session. During this session, measurements followed one another in a fairly rapid sequence precisely as they had done in Judd's investigation. Thus, since no question of different schedules arises, fixation is the only factor which can have caused the very high speed with which the illusion was destroyed in these cases. As an example, we give

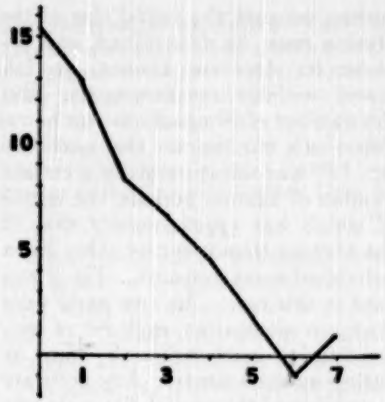


FIG. 3. Results during the first experimental session of subject B: the size of the illusion in mm. as a function of the growing number of trials

the curve which represents results during the first session of Subject B (Fig. 3).[8] To a degree, the same factor may have lowered the figures of the three other *Ss*; but we are unable to say how strong this influence has been. For, with *Ss* C–E, the illusion did not disappear until several sessions had been given which were more widely distributed in time than sessions had been in earlier work. Such a distribution, it will later be seen, has the same effect upon the decrease of the illusion as fixation seems to have.

According to K. and W., the size of the illusion soon decreases when *Ss* only *fixate* a point of the pattern. One may conclude from this observation that a sequence of mere fixation periods will destroy the illusion entirely. If this were the case, the satiation theory of the effect would be greatly strengthened. Naturally, comparisons cannot be avoided alto-

[8] The points of the curve are averages of four successive trials. Obviously, the development would not be adequately represented if averages have been computed from larger numbers of individual measurements.

gether, because the initial size of the illusion must be determined, and because its decrease cannot be followed without measurements. But the number of comparisons can be reduced to a minimum. Our next subject (F) was therefore given a certain number of fixation periods, the length of which was approximately that of the average time spent by other *S*s on individual measurements. Fig. 2 was used in this case. Its two parts were shown in alternation, each for 12 sec., the fixation mark being the same as during measurements. Any accurate comparing of the crucial distances was made impossible by separating successive fixation periods by intervals of 48 sec. For good reasons, it was assumed that in spite of these intervals satiation of a given part would to a degree survive from one of its presentations to the next.

The initial size of the illusion was found to be 18 percent. After a sequence of 24 inspection periods (12 for each part of the pattern), the average of four measurements was 9.5 percent, i.e., the illusion was now reduced to one-half of its original size. Only 16 inspection periods (8 for each part) were now given. After their completion, four measurements showed no evidence of an illusion. At the time, the parts of the pattern had been inspected 20 times, i.e., each for 4 min. altogether, and these parts had been compared 8 times. Just as in the case of *S*s A and B, the process was completed in less than an hour. The suggestion that the illusion was destroyed during the 8 comparisons, and that therefore learning might have destroyed it, is surely not compatible with the evidence which one finds in previous investigations. We agree that, in spite of their small number, the comparisons may have contributed to the final outcome, if merely because satiation probably grew while they took place. It seems to us, however, that the experiment as a whole can satisfactorily be explained only if the sequence of inspection periods is regarded as the main cause of the effect. Thus, since during these periods *S* cannot have learned to compare better, an interpretation of the effect in terms of satiation seems to us indicated.

With two other *S*s, we used only the upper part of Fig. 2, and replaced the lower part by two short vertical lines, the left one of which could be moved until its distance from the second line appeared equal to the distance between the corners of the upper part. Repeated inspection periods were again given, and the size of the illusion was only occasionally measured in a few trials. Unfortunately, the illusion which was inconspicuous in both cases disappeared so fast that both the number of inspection periods and that of comparisons were too small to permit any conclusion as to the relative influence of the two factors.

If the illusion were destroyed because *S*s learn to compare better, then the process ought to come to an end when comparisons can no longer be improved, in other words, when the illusion has completely disappeared. But if the explanation in terms of satiation is accepted, the process may continue beyond this point until the effect of satiation is as strong as it can become under the given conditions. There is no particular reason why this limit should coincide with the point of objective equality. With three *S*s, experimentation was therefore not interrupted when the illusion had disappeared. As a consequence, these *S*s developed *negative* illusions. The part of the pattern which normally looks smaller now appeared larger. In all three cases experimentation was discontinued when the size of the

illusion was −1 cm., or −14 percent. Subject A may serve as an example. After her first session, in which the illusion had been destroyed, this *S* was given shorter sessions for several days. Altogether, nine averages were now determined from groups of four individual measurements each. Without exception, these averages were negative, the last three being −.8, −.9 and −1 cm. At this point, the negative illusion was greater than the positive illusion had been in the beginning (+.55 cm.). Among the *Ss* with whom experiments were continued when the illusion had disappeared was F, who so far had been given mainly inspection periods rather than measurements. Under the same regime, this *S* also developed a negative illusion. Its size reached the value −1 cm. when the eight preceding averages had been negative.

We do not feel that further new evidence is needed in this respect, because Judd and Seashore mention the same fact. With one of Judd's *Ss*, 'some negative settings' occurred at the time when the illusion had just disappeared. Of Seashore's four *Ss*, two eventually developed negative illusions. The authors offer no explanation. Since, at the time, learning appeared as the only process which could cause the destruction of the illusion, they must have regarded such observations as mere curiosities. At present, it seems more satisfactory to conclude that, in their experiments just as in ours, the effect of satiation finally became more effective than the factors which cause the illusion.

When the *Ss* of Cameron and Steele (2) had made about 2000 experiments with the Poggendorf pattern, negative settings were also found; even the average of 25 measurements was now occasionally negative. The same happened after many trials when Judd and Courten (6) investigated the disappearance of the Zöllner illusion. If we are right in assuming that, with the M.-L. pattern, negative illusions are caused by extreme degrees of counteracting satiation, then the negative values obtained with the other illusions are probably established in the same fashion. This would imply, however, that not only the M.-L. phenomenon but also the other illusions are destroyed by satiation rather than by learning.

From the preceding argument we can derive a further prediction. If, with some *Ss*, satiation is so intense that eventually its effect becomes stronger than the factors which cause the illusion, with other *Ss* the process may be so weak that the illusion can never be entirely destroyed. From the work of K. and W. it is known that the size of figural after-effects varies widely from one *S* to another. *Ss* with extremely small sizes of these effects seem to be comparatively rare. But there are such persons. Thus, if it is actually satiation which causes the illusion to decrease, its complete destruction may not be possible in such instances.

In this connection, our results with four further *Ss* (G–J) deserve to be mentioned. Their experiments were interrupted when the illusion had not yet disappeared.

With G, whose initial measurements had given the value 39 percent, 166 individual trials reduced the illusion to 17 percent. At the time, it was still slowly decreasing, and might well have disappeared if further experimentation had been possible.

With H, the initial amount 27 percent decreased to 10 percent under the influence of 64 measurements. When 48 additional measurements caused no further change, experimentation with this *S* was discontinued.

With I, the size of the illusion dropped from 31 to 11 percent within only 44 individual measurements. But the next 100 trials had no effect

whatsowever, although their distribution in time was favorable. Further experimentation seemed therefore useless.

We cannot, of course, say what would actually have happened if *Ss* H and I had been given 500 or 1000 further trials. But in both cases the transition from a fairly rapid decrease of the illusion during early trials to an apparently constant condition afterwards was so striking that some essential limit had probably been reached when the transition occurred. We regard it as a plausible thesis that at the crucial stage the maximum of satiation possible with these *Ss* had been established.

Observations with J greatly strengthened our suspicion that it is by no means possible to destroy the M.-L. illusion with all *Ss*. For, with this *S*, repeated trials did not even affect its initial size. This size was found to be 25 percent, and the average of 168 further measurements was 26 percent. Results were so constant that, from this mean, the means of 42 groups of four individual measurements each deviated, on the average, by less than one percent of the length of the crucial distance. And yet, in this instance we tried all variations of the experimental schedule which had proved to accelerate the process with other *Ss*. Because of observations such as these, our statements about the accelerating influence of fixation on the disappearance of the illusion had to be formulated with a certain reserve. There seem to be instances in which no development occurs which could be accelerated. It should perhaps be mentioned that this particular *S* was an excellent student of Swarthmore College, most eager to work for us, and apparently an altogether 'normal' person. Fortunately, this *S* had not yet become acquainted with the concept of figural after-effects or with any observations in this field. This enabled us to make a number of tests, the meaning of which was unknown to her. We found that, with most combinations of inspection- and test-objects, this *S* saw no figural after-effects at all. Only when situations were used which give the most striking results with other *Ss* were weak effects at last established. If, in a large group of *Ss*, the speed with which the M.-L. illusion disappears were compared with the size of their figural after-effects, the two sets of measurements might show a reliable correlation.

It happens that at this point our own results are again confirmed by an earlier investigation. One of Seashore's *Ss* made measurements for 24 days. Experiments were discontinued when at the end of this period the illusion was just as great as it had been in the beginning.[7]

K. and W. describe an observation in which, after prolonged inspection of the M.-L. pattern, three short vertical lines were shown which passed through the points previously occupied by the corners of the angles. In relation to these points, the fixation mark was the same as it had been during the inspection period. Under these circumstances, the distance which is visually shorter in the illusion pattern appeared too long (**8**, p. 349 f.). The observation demonstrates the fact that in the M.-L. figure satiation is asymmetrically distributed: the two parallel lines on the side which looks shorter in the illusion

[7] The theoretical interpretation of such results is perhaps less simple than it might at first appear. One feels inclined to say that in *Ss* such as J satiation must be exceptionally weak, or that it must fade particularly fast. But if we are not mistaken, one could also make the opposite assumption. There can be little doubt that, as a consequence of everyday vision, the visual cortex as a whole is in a state of quasi-permanent satiation, and that local satiation as established in special fixation periods is merely added to the level of that persistent condition (cf. K. and W., **8**, pp. 341–346). Now, satiation is a self-limiting process (cf. **8**, p. 329). It therefore seems at least possible that, with some *Ss*, temporary and local satiation effects are found to be weak because the more persistent satiation of the tissue as a whole is already unusually strong.

lie just outside a highly satiated region, and their distance must grow. The opposite must happen on the other side where the lines lie just inside the most affected areas. Both changes together must make the objectively equal distances between the three lines appear unequal in the sense which was actually observed.

As a further corroboration of our thesis that the M.-L. illusion is weakened or destroyed by satiation, we repeated this experiment with six *Ss* when their M.-L. illusion had disappeared. The vertical line on the left side could be shifted, and the amount of the distortion could therefore be measured. In all cases, the qualitative observation of K. and W. was confirmed. When the illusion had disappeared, the distance corresponding to the part of the M.-L. figure, which looks smaller in the illusion, was clearly greater in the test pattern. With *Ss* A–F respectively, this distance was now too long by 12, 11, 6, 14,6, and 12 percent. These figures prove that, after prolonged experimentation with the M.-L. pattern, unequal amounts of satiation are actually established in its two parts, and that the direction of the asymmetry is such as to operate against the illusion. Since in this proof a test pattern is being used from which the main characteristics of the illusion pattern are absent, the proof is independent of the issue under discussion, namely, the decrease of the illusion as such.

Similar tests might be used when attempts are made to answer a question which we have raised in the beginning. If the interpretation of the 'practice' effect in terms of satiation is to be generally applicable, it must be postulated that about the same pattern of satiation tends to develop without fixation as has just been shown to exist after trials with fixation. Consequently, when the illusion has been destroyed under the former condition, two distances coincident with those of the M.-L. pattern but limited by short vertical lines should also appear unequal in the sense just indicated.

Although experiments of this kind have not yet been performed, one feels tempted to predict their result on the basis of facts which are already known. It seems quite possible that Judd and Lewis themselves would have hesitated to accept the 'practice' theory, if at the time any other theoretical possibility had been available. Occasionally, they must have been disturbed by their own findings.[3] Thus their observations prove that, when the illusion has been destroyed, it is absent only in the particular area in which the pattern has been shown during the preceding trials. When Judd and Courten investigated the disappearance of the Zöllner illusion, it was ". . . after practice still possible to see the illusion by changing one's position. Thus getting up from in front of the figure and walking away to a new position or changing the position of the head induced the illusion as before practice" (6, p. 123). The same observation was made by Cameron and Steele (2) when the Poggendorff illusion had been destroyed. It appears to us improbable that the factors responsible for the disappearance of these illusions differ from those which destroy the M.-L. illusion. Similar observations are therefore to be expected in the latter case as well. Actually, one of our own *Ss* once spontaneously remarked that a

[3] Both authors express their views with considerable caution. Judd, for instance, who was particularly interested in the rôle of eye movements, might easily have been misled by his observation that, when experiments were often repeated, the size of eye movements along the parts of the illusion figure tended to change in the direction in which the appearance of these parts changed. And yet, both he and Lewis argue against an interpretation of the 'practice' effect in terms of eye movements or corresponding kinesthetic sensations. From our present point of view, Judd's observation indicates that the amplitude of eye movements across an object varies when the size of this object is changed by satiation.

setting, in which the two distances had just appeared equal, again showed a strong illusion when he moved to a greater distance.

This *S*, of course, fixated while the experiments were done. But Judd and Lewis have proved in special experiments that the so-called 'practice' effect is localized also under the conditions of their own work with the M.-L. pattern. After destruction of the illusion with a given arrangement of the pattern, it appeared again, sometimes in almost its full size, sometimes more clearly smaller, but once even enhanced, when the length of the crucial distances was altered, when the pattern was turned by 90°, or when each part was placed in the position which the other part had occupied before. Observations of this kind obviously support the thesis that repeated trials destroy the illusion by means of satiation. On the other hand, we fail to see how such facts can be reconciled with the view that the illusion is destroyed by learning. Within very wide limits, the recognition of figures, the recall of items associated with visual objects, and learned ways of dealing with patterns, say, in geometry, are known hardly to be affected by displacements in space, by changes of size, and so forth. Hence, at this point, the thesis that the illusion is destroyed by learning actually appears to be at odds with well-known evidence in the field of learning itself. For example, none of the learned performances to which we have just referred would be appreciably disturbed, if an *S* merely looked at the objects in question from a somewhat greater distance. But, in doing so, he would, of course, change the spatial relation between such objects and satiation patterns, if such patterns had previously been established.

Although the 'practice' effect is fairly well localized, there is one point in these observarions that deserves more particular consideration. Both Judd and Lewis found that, when the pattern was shown in a new position, the illusion had, for the most part, not its full size, and could now be destroyed in fewer trials. Are such observations explainable on the premise that the illusion is destroyed by satiation? Once before, we mentioned that the initial size of the illusion as measured by Judd and Lewis is likely to be too small, because they derive this size from large groups of trials, in the course of which a certain degree of satiation must already be established. Since this argument applies to first experiments both in the original and in the changed position of the pattern, it cannot alone explain the fact that in a new position the illusion has not its full size. But when the argument is combined with a further principle, a plausible explanation can be given. To make this clear, we must for a moment turn to the theory of satiation as such.

According to K. and W., satiation is merely another name for the biophysical concept of electrotonus. The authors assume that, when parts of the visual field are segregated as objects or figures, corresponding cortical areas are pervaded and surrounded by direct currents.[9] But currents which spread in living tissue tend to affect this medium by electrotonic action, which means that they polarize the surfaces of cells through which the flow passes, and that, presently, the polarizability of all anodic surfaces is also increased.[10] Currents which originate in the nervous system have the same effect upon their substratum. Quite apart from other consequences, electrotonus constitutes an obstacle to

[9] Recently, this hypothesis has been supported by the observation that under conditions of pattern vision smooth potentials can be registered from the human head (7).

[10] In a recent article we find the statement that, in their explanation of figural after-effects, K. and W. refer to an 'hypothetical electrotonus.' The concept, it seems, is regarded as a product of contemporary speculation. Actually the term electrotonus denotes a group of well-known facts in electrophysiology, and was thus used by Du Bois-Reymond, who discovered these facts in 1843.

the flow by which it is caused; in other words, this current is weakened in regions in which it was first most intense, and therefore caused a maximum of electrotonus, and it is strengthened in regions where it was first weaker, and thus affected the medium less. In the theory proposed by K. and W., such distortions of figure currents by patterns of satiation are made responsible for figural after-effects. The currents which suffer distortion may either be those of the figures also shown during the satiation period; in this case, the figures will gradually alter their own appearance. Or different figures are shown as test objects in the affected region; then these objects will be altered by the obstruction.

What must happen if, with a given figure, a pattern of satiation is first established, in which certain areas are much more strongly affected than others, and if then a new figure causes further satiation in the same region? The distribution of the current of the new object must differ from the flow which would normally be associated with this object. As we have just seen, this current will be weakened in the areas which have been most affected by the current of the first figure, and it will be intensified in areas which have previously been less affected. It follows that satiation by the new current must develop in a similarly 'abnormal' fashion: in the former areas, the new satiation will be weak, but in places which have not been strongly satiated by the current of the first figure the new object must satiate the tissue more rapidly than it would if the first satiation had not occurred. In a shorter formulation, once an uneven pattern of satiation is established, further satiation will always be accelerated in parts in which it has so far been less intense.

This reasoning applies directly to experiments in which, after destruction of the M.-L. illusion, the pattern is repeatedly shown in a new position. As an example, we may consider the case in which each part of the pattern now occupies the place in which the other part was shown during the 'practice' period (cf. Fig. 4, in which the new position of the pattern is given in continuous lines, the old position in broken lines, and the areas of maximal previous satiation are marked S.). In this situation, a strongly satiated region lies between the two angles on the right side of the new pattern. Consequently, the new current is intensified in the interior of these angles, which will now be satiated with abnormal speed. Similarly, the areas of the two new angles on the left side will now be pervaded by an unusually intense flow, and will therefore suffer correspondingly rapid satiation. As a result, the right

FIG. 4. Relation of the reversed Müller-Lyer figure to the areas in which maximal satiation has been established before the reversal

side of the new pattern will shrink with unusual speed, and the expansion of the left side will also be greatly accelerated. Conditions for the destruction of the illusion in repeated trials are therefore now particularly favorable.

The present reasoning explains the fact that when Judd and Lewis turned the pattern in space repeated trials appeared to be more effective than they had been in its original position. The same explanation may be applied to the observation that in the new position the initial size of the illusion was generally found to be smaller. For, if measurements of this size always give too small a value when many trials were used for the purpose, then the initial size must be further reduced when the process of satiation is accelerated. It seems that this conclusion can be subjected to a simple test. If, in its new position, the pattern were shown only for a very short time, a stronger rather than a weaker illusion ought to be observed. This follows from the very fact which, in repeated trials, must cause the illusion to decrease with unusual speed: the satiation which has been established during trials in the original position of the pattern intensifies the flow in the interior

of the new angles. According to K. and W., this distortion of the new currents must make the right distance in Fig. 4 appear larger, and the left side shorter, than they are in the 'normal' illusion. Only when inspection is prolonged, or when trials are repeated, can this first effect be overcome by the accelerated satiation which the distorted flow must now cause.[11]

Lewis made experiments not only under ordinary conditions of observation but also when the pattern was shown tachistoscopically. In the latter case, the time of exposure was less than .02 sec.[12] Several questions arise from the results of the tachistoscopic tests. In the *first* place, the initial size of the illusion was clearly greater under these conditions than it was when *S*s were given more time for the inspection of the pattern. General difficulties involved in tachistoscopic perception cannot explain the difference, because actually *S*s soon judged with greater confidence and more consistency in the tachistoscopic situation.[13] As another explanation, it might be suggested that, during tachistoscopic exposures, *S*s have no time to compare the crucial distances in an analytical attitude, and that for this reason the illusion must now be greater. This may be ture. But it may also be pointed out that if, under ordinary circumstances, satiation reduces the initial size of the illusion when this size is derived from many individual comparisons, then very short tachistoscopic trials must give a larger initial size, simply because they cause much less satiation.

This explanation agrees with a *second* result of the tachistoscopic experiments. While, with Lewis' *S*s, the illusion always disappeared when ordinary observation was used in trial after tiral, it was, if at all, only slightly affected when the pattern was tachistoscopically presented. Again, it might be said that the latter condition makes it impossible for *S*s to acquire an analytical attitude, and that the illusion does not disappear for this reason. But it is also obvious that if, under ordinary conditions, satiation destroys the illusion, its influence must be greatly reduced when exposures are restricted to less than .02 sec. To be sure, K. and W. have shown by a summation technique that reduction of exposures to .25 sec. does not prevent satiation. But it does not follow that, after a further reduction at the ratio 12:1, satiation must still be strong or persistent enough to survive from one trial to the next. Fortunately, this issue can be subjected to simple experimental tests. Such tests may also serve to relate psychological experimentation in this field to known facts about the temporal characteristics of electrontonus.

Lewis does not explicitly state a *third* result of his tachistoscopic experiments. But in view of his experimental schedule and the information actually given, there can be no doubt about this result as such. The same *S*s participated in the measurements under ordinary conditions of observation and in the tachistoscopic trials. Moreover, experiments of both kinds were done on the same days.[14] It follows that, at the end of their trials, these *S*s saw no illusion when the pattern could be inspected for some time, while they saw the illusion in approximately its original size when exposures were restricted to less than .02 sec. On the assumption that the illusion is destroyed because *S*s learn to concentrate more specifically on the crucial distances, this

[11] In one of Judd's experiments, the illusion was actually enhanced in the new position of the pattern. It is interesting to note that in this case the size of the illusion in the new position was derived from only 10 individual measurements.

[12] We find this figure in an article by Hicks and Rivers (3), whose apparatus and method were used by Lewis.

[13] If the satiation theory is accepted, this observation is by no means surprising. When *S*s have sufficient time to do so, they will usually try to make sure that their first impression is reliable, and will therefore postpone their final decision. Since, during this period, satiation can change the appearance of the pattern, they may find themselves disturbed by contradictory impressions. Under the same circumstances, eye movements may, of course, further complicate the situation. The tachistoscopic procedure eliminates these causes of uncertainty.

[14] Cf. 3, p. 247.

curious fact can perhaps be explained as follows. Even when, in trials under ordinary conditions, *Ss* have learned to assume a highly analytical attitude, and thus to overcome the illusion, actual analysis remains a process which takes a certain time. But sufficient time is given only so long as the duration of the exposures is not limited; it is not given in the tachistoscopic trials. Hence, *Ss* cannot apply to the latter trials what they have learned in the former, and Lewis' result becomes necessary.

Can this result also be explained on the premise that satiation rather than learning destroys the illusion? At first, no such explanation may seem to be possible. It is quite true that very short tachistoscopic exposures may not be able to cause sufficiently persistent and therefore gradually growing satiation. But if experiments under ordinary conditions are performed at the same time, and if during these experiments high degrees of satiation are really established, why should this condition become ineffective when the time of the exposures is restricted? This is indeed a crucial question. We will now show that a simple answer follows from the theory of satiation and of figural after-effects. Electrotonic action has two components: first, actual polarization of cell surfaces, which is a process of relatively short duration, and secondly, changes of anodic surfaces by which their polariazbility is gradually increased. Actual polarization is likely to disappear very soon after the current by which it is caused has been discontinued. For this reason, only the other component, which tends to persist for long periods, can be identified with the enduring affection of the tissue which we call satiation. Nevertheless, if we follow the analysis given by K. and W. (8, pp. 320 ff.), the fact that in a certain region the polarizability of cell surfaces is increased, i.e., that the region is satiated, cannot as such distort figure currents, and thus cause figural after-effects. Such effects can arise only when now the affected region is actually polarized to the degree which its increased polarizability makes possible. Obviously, when figural after-effects are being observed, the current which polarizes the satiated area is that of the test object. It follows that in all figural after-effects this object must play an active rôle. For it is its current which establishes the obstructing and distorting forces that can now be set up in the satiated region. Once this is realized, we have no difficulty in explaining why, in Lewis' experiments, the intense satiation caused under ordinary conditions of observation had almost no effect upon measurements made under tachistoscopic conditions. When explaining why the illusion was not destroyed in repeated tachistoscopic trials, we said that during exposures which are shorter than .02 sec. satiation is probably not intense enough for the purpose. We now have to add merely that under the same conditions the duration of the figure current is probably not long enough to cause a sufficient degree of actual polarization, and that therefore the illusion is seen in approximately its full size, although the polarizability of the crucial areas has been greatly increased during the trials with longer exposures.

It will be realized that the principle which has just been formulated must hold quite generally. If it applies to Lewis' finding, then it must also apply to any figural after-effects whatsoever. In other words, however great given figural after-effects may be when measured in the usual fashion, their size ought to be considerably reduced when test objects are given in very short exposures. It will thus be pos-

sible to test a further consqeuence of the theory of figural after-effects.[15]

This consequence deserves particular attention because, if it were verified, it might establish an interesting relation between the study of figural after-effects and recent observations in brain pathology. Bender and Teuber (1) report that, in the visual field of patients in whom parts of the occipital brain have been injured, objects are displaced from their normal positions. When the objects are shown in the neighborhood of the injured area, their distance from this area is increased. The displacement can therefore be said to have the same direction as it has in figural after-effects, if the injured part of the tissue is regarded as comparable to a highly satiated part of normal tissue. The authors themselves mention that their observations resemble figural after-effects. One feels tempted to infer that, quite apart from any strictly local disturbance of function, the injured tissue of such patients represents a highly polarizable area, that this area is polarized by the currents of test objects in the neighborhood, and that, as a consequence, such objects recede from the place of the defect. We do not, of course, know whether the polarizability of brain tissue actually grows when certain pathological conditions are established. There is, however, a further reason for considering this possibility. When the patients were examined tachistoscopically rather than with longer exposures, the test objects were not displaced—precisely as, under the same conditions, the M.-L. illusion does not seem to be affected by a previously established satiation. Thus, if our explanation of the latter fact should prove to be correct, this explanation could also be applied to the result of tachistoscopic tests in the work of Bender and Teuber. But, to repeat, this reasoning depends on the assumption that the injured brain tissue of their patients was abnormally polarizable.

[15] A preliminary experiment of this kind was recently done by Mr. W. Lichten of Swarthmore College. He used a pattern which involves no illusion. When a highly persistent figural after-effect had been established, this effect was tested also under tachistoscopic conditions. The exposure time was not quite so short as the one used by Lewis. Nevertheless, the after-effect proved to be reliably smaller when measured in this situation. More experiments will have to be performed, because so far only one *S* was tested.

Summary and Conclusions

Experimental evidence suggests that repeated measurements destroy the Müller-Lyer illusion by satiation rather than by learning. The facts which support this thesis are as follows:

1. The illusion can be destroyed when *S*s do not know that they are dealing with an illusion pattern, and when they are not told how their results compare with objective facts.

2. The illusion can be destroyed when measurements are largely replaced by satiation periods.

3. When after destruction of the illusion experiments are continued, negative illusions of considerable size may be established.

4. There is some evidence that the illusion cannot be completely destroyed with all *S*s.

5. When after destruction of the illusion a 'normal' test pattern is presented in the area of the illusion figure, this pattern exhibits the after-effect which is to be expected if the illusion is destroyed by satiation.

6. When the illusion has been destroyed in one position, it is still present in other positions.

7. The fact that the illusion can be more rapidly destroyed in new posi-

tions has been derived from the theory of satiation.

8. The observation that the illusion cannot be destroyed in tachistoscopic experiments, and that after its destruction under ordinary conditions it hardly appears affected in tachistoscopic tests, follows from the theory of figural after-effects. This theory seems to relate recent findings in brain pathology to facts in normal perception.

9. As a general rule, experiments in which given situations are presented many times in succession can be made only with considerable caution, because under such conditions satiation may affect the situations under investigation.

(Manuscript received for immediate publication January 19, 1950)

References

1. Bender, M. B., & Teuber, H. L. Spatial organization of visual perception following injury to the brain. *Arch. Neurol. Psychiat.*, 1947, **58**, 721–739; and 1948, **59**, 39–62.
2. Cameron, E. H., & Steele, W. M. The Poggendorff illusion. *Psychol. Rev., Mon. Suppl.*, 1905, **7**, 83–111.
3. Hicks, G. D., & Rivers, W. H. R. The illusion of compared horizontal and vertical lines. *Brit. J. Psychol.*, 1908, **2**, 243–260.
4. Judd, C. H. Practice and its effects on the perception of illusions. *Psychol. Rev.*, 1902, **9**, 27–39.
5. Judd, C. H. The Müller-Lyer illusion. *Psychol. Rev., Mon. Suppl.*, 1905, **7**, 55–81.
6. Judd, C. H., & Courten, H. C. The Zöllner illusion. *Psychol. Rev., Mon. Suppl.*, 1905, **7**, 112–139.
7. Köhler, W., & Held, R. The cortical correlate of pattern vision. *Science*, 1949, **110**, 414–419.
8. Köhler, W., & Wallach, H. Figural after-effects. An investigation of visual processes. *Proc. Amer. Philos. Soc.*, 1944, **88**, 269–357.
9. Lewis, E. O. The effect of practice on the perception of the Müller-Lyer illusion. *Brit. J. Psychol.*, 1908, **2**, 294–306.
10. Seashore, C. E., *et al.* The effect of practice on normal illusions. *Psychol. Rev., Mon. Suppl.*, 1908, **9**, 103–148.
11. Woodworth, R. S. *Experimental psychology.* New York: Holt & Co., 1938.

A COMPARISON OF CRITICAL FLICKER FREQUENCIES UNDER CONDITIONS OF MONOCULAR AND BINOCULAR STIMULATION *

BY FRED H. IRELAND

Fordham University

Introduction

The problem of central versus peripheral control of various visual functions has long been a controversial subject, and numerous experiments have been conducted in an effort to clarify the issue.

The present investigation constitutes a repetition of an earlier experiment with intermittent light stimuli performed by Sherrington (7) in 1904. His purpose was to obtain clues to the limiting mechanism responsible for critical flicker frequency, i.e., that frequency above which a pulsating light source is experienced as a continuous stimulus of uniform brightness. Sherrington was interested in determining whether differences existed among CFF's obtained under various conditions of monocular and binocular stimulation, since such differences would indicate an interaction between the two eyes.

In order to produce flicker of the desired frequency Sherrington employed a system of perforated cylinders rotated about a concentrically located light source. The speed of the electric driving motor was controlled by a rheostat. An optical viewing system permitted monocular stimulation of either eye, or stimulation of conjugate retinal areas of both eyes by the same light source.

The two conditions of binocular stimulation which subsequently received particular attention were: (a) in-phase flicker with synchronous stimulation of both eyes and (b) out-of-phase flicker, where both eyes were stimulated by the same light.

In comparing the CFF's obtained under those conditions of stimulation, Sherrington stated that: "As far as sensual effect goes, the light phases at the one eye practically do not, therefore, interfere or combine at all with the coincident dark phases at the other; and conversely. Nor do they, in the alternate left and right arrangement, add themselves as a series of additional stimuli to the like series of stimuli applied at the other eye" (7, p. 37). Hence he concluded that it was the peripheral mechanism of the retina which was responsible for the effective fusion of successive light impulses above a certain frequency. Many subsequent investigators in the field of vision have placed great faith in these findings, although, to date, no confirmatory studies have been undertaken (2, p. 1049; 8).

However, Bartley (1) and others, with such phenomena as binocular color mixture and Fechner's paradox in mind, voice a note of caution against blindly accepting the theory of exclusive peripheral determination and suggest that further investigation is necessary.

Sherrington gives no exact information concerning the number of subjects tested in the original experiment, although a table pertaining to the experiment (7, p. 35) seems to indicate that perhaps only four individuals were used. The degree of frequency stability of the mechanical device employed to produce the stimulus is also not indicated. Usually the precision obtainable with such instruments does not equal that of modern electronic equipment. Furthermore, auditory cues arising from the electric motor and driving mechanism are difficult to eliminate.

In view of the above considerations, it appeared desirable to repeat Sherrington's experiment with such modi-

* The writer is indebted to Dr. Robert T. Rock, Jr. and Dr. Henryk Misiak for their helpful advice and suggestions throughout the experiment.

fications of procedure and equipment as would be provided by more modern methods and instruments.

Apparatus

In order to minimize instrument error, a high precision electronic flicker apparatus was developed by the author. Since this apparatus has already been described in detail in a separate article (4), only a brief mention will be made here of its functional characteristics.

Instead of the single light source employed in Sherrington's experiment, two gaseous crater lamps were here used as stimuli. These were located in a stereoscopic viewing device in order to provide binocular stimulation of conjugate retinal areas. Either light could be blocked from view by a metal shield during monocular stimulation. In this optical arrangement the visual angle subtended by each test patch was 2°22′. An electronic pulse generator was used to energize the two lamps. The flicker frequency could be varied instantaneously and continuously over the total range of from 20 to 60 cycles per sec. Frequency calibration could be checked against the 60 c.p.s. house-current at any time during the experiment and was maintained to a tolerance of within one-tenth of one percent of the indicated dial readings. By changing the phase relationship of the impulses applied to the two stimulus lamps, these could be made to flash either synchronously (in phase) or out-of-phase with one another by exactly 180°. On the screen of a cathode-ray monitor the on-off ratio, the wave shape and the phase relationship were easily observable. Parameters of the electronic control circuits were adjusted to yield light pulses of rectangular shape. Main points of consideration were steep initial and terminal slopes and an on-off ratio of 1:1. The brightness of the test patch as measured through the ocular lens system was 0.44 millilambert. The same level of illumination was maintained throughout the experiment. A keying circuit enabled the experimenter to initiate and terminate the stimuli at will.

The investigation was carried out in a small vision laboratory. In order to eliminate light reflections the walls and ceiling of this room had been painted black. A wooden partition separated the subject and the experimenter. The modified stereoscope, secured to the top of a small table, was adjusted to eye level of the seated subject. All remaining instruments (pulse generator, monitoring equipment etc.), arranged for convenient manipulation, were located on the other side of the partition upon a second table.

Methods and Procedure

Twenty-four subjects, 20 males and four females, were tested. Their ages ranged from 20 to 40 years, with the majority in the 20 to 25 year age group.

During preliminary examinations the subjects' eye dominance was established by means of the 'Miles A-B-C Vision Test for Ocular Dominance.' Also, because of the importance of stimulating corresponding retinal areas under conditions of binocular flicker, all subjects were tested for normal stereoscopic vision.

Four conditions of stimulation were employed: (a) monocular flicker—dominant eye; (b) monocular flicker—non-dominant eye; (c) binocular flicker—out-of-phase; and (d) binocular flicker—in-phase.

To compensate for possible interaction of these stimuli the sequence of conditions was changed for each individual. Every subject was tested twice for each condition of stimulation on two occasions separated from two to five days. The order of presentation of stimuli for the subject's second sitting was the reverse of that followed during the first session.

The method of limits was used to obtain the CFF's. Ten determinations were made to establish the CFF for a particular condition of stimulation. Five readings were taken in passing from the state of flicker to one of apparently steady light, while the remaining five readings were obtained on a descending frequency scale going from a state of fusion to one in which the subject again experienced flicker.

The testing session lasted about 25 min. After remaining in darkness for five min. to adapt the foveal region of the eye to the low level of illumination emanating from the test patches, the subject was asked to place his head against the visor of the instrument. He was then acquainted with the conditions of 'flicker' and 'no flicker' by observing first a stimulus of extremely low frequency (20 c.p.s.) and then one of higher frequency (60 c.p.s.).

For the actual determination of a CFF the dial was set to the lower end of the scale and the stimulus key was thrown into the 'on' position. When the subject indicated that he perceived flicker, the experimenter moved the dial at a relatively constant rate of speed toward the higher frequency end of the band until the subject indicated (by saying "now") that the point of fusion had been reached. The stimulus was then interrupted for a period of seven seconds during which the subject could rest his eyes while the experimenter recorded the dial readings. Next the dial was set off in the high frequency direction and the procedure was repeated on a descending frequency scale. Thus, 10 readings

were obtained for each of the four conditions of stimulation for every subject.

Results and Discussion

Dial readings were first converted into corresponding frequencies by means of the calibration curve for the instrument. A mean CFF value and standard deviation for each distribution of 10 readings were computed. As Table I indicates, a positive frequency shift had occurred from Session I to Session II for all individuals.

TABLE I

Mean CFF's Obtained for the Various Conditions of Stimulation of the 24 Subjects During the Two Testing Sessions

Condition of Stimulation	Mean CFF's	
	Session I	Session II
Monocular:		
dominant eye	29.8	31.2
non-dominant eye	29.9	31.0
Binocular:		
in-phase	31.5	32.8
out-of-phase	28.5	29.4

Hence the results from each session were treated separately. The tendency for CFF's to be higher upon retest was also observed by Misiak (5) in his experimental studies of flicker.

Mean standard deviations, computed as measures of intra-individual variability for the different conditions, were found to range from 0.57 to 1.01 c.p.s. with an average value of 0.71.

In order to compute the consistency of measurement of each CFF during both testing sessions, reliability coefficients were computed by the split-half method of correlation. Thus, for each individual two means were computed for every set of readings obtained under a particular condition of stimulation. One mean was determined for the first, second, fifth, and sixth readings, and another for the third, fourth, seventh, and eighth readings. The two sets of 24 means thus obtained for a given condition of stimulation were then correlated. Finally, the reliability for the full set of 10 readings was estimated by means of the Spearman-Brown prophecy formula with N equated to 2.5. For all conditions of stimulation these estimated reliability coefficients were above .99.

The mean CFF for monocular stimulation (stimulus brightness = .44 millilamberts, on-off ratio = 1:1) in this experiment closely agrees with the value obtained from the curve plotted by Crozier and Wolf (3).

An inspection of Table I will show that differences exist among the CFF's

TABLE II

Evaluation of the Differences in Mean CFF's Obtained for the Various Conditions of Stimulation During the Two Testing Sessions. (Numbers Given in Parentheses are the Mean Frequencies Obtained for the Indicated Condition of Stimulation.)

Compared Conditions		Differences	t-values
Session I			
dominant non-dominant	(29.8) (29.9)	.1	.6
dominant in-phase	(29.8) (31.5)	1.7	8.8*
dominant out-of-phase	(29.8) (28.5)	1.3	5.4*
Session II			
dominant non-dominant	(31.2) (31.0)	.2	1.7
dominant in-phase	(31.2) (32.8)	1.6	8.8*
dominant out-of-phase	(31.2) (29.4)	1.8	8.4*

* Significant at the .01 level of confidence.

obtained under the various conditions of stimulation. These differences were evaluated for their statistical significance by determining the *t*-ratios which are presented in Table II, together with a listing of the compared conditions. Statistically significant values are indicated by asterisks. Because the differences between CFF's obtained for dominant and non-dominant eye stimulation proved to be statistically insignificant in both testing sessions, further comparisons of binocular CFF's were made only with those values obtained for the dominant eye.

Table II indicates that statistically highly significant differences in CFF were obtained when results for binocular stimulation were compared with those found when just one eye was stimulated. Synchronous binocular stimulation resulted in an elevated CFF, while out-of-phase flicker produced a lowering of the fusion frequency. Sherrington, under similar stimulus conditions, also found very slight shifts of CFF in the same directions. However, since there were only minute variations which were not consistent from one individual to the next, Sherrington considered them insignificant. He states that, "the broad outcome of the above experimental observations is that so far from bright phases at the one eye effacing dark phases at the corresponding spot of the other eye, there is hardly the merest trace of any such interference" (7, p. 37).

The finding that synchronous stimulation of the retinas yields a significantly elevated CFF agrees with the results obtained by Misiak (6), who obtained higher CFF's when two eyes were stimulated by the same flickering test patch. These results seem to indicate a certain degree of binocular summation, whereas the lower fusion frequencies found in binocular out-of-phase stimulation tend to support a theory of inhibitory interaction beteen the two eyes. It is possible that in the latter case the condition of Fechner's paradox prevails wherein, for a certain range of light intensities, the sensation resulting from binocular stimulation with steady light is equal to the average value of the monocular sensations. Thus, in the present experiment, the effective light intensity of the stimulus may have been reduced to a value intermediate between zero stimulation of one eye and maximum intensity at the other. This may account for the lowered CFF, for it is known that the frequency of fusion varies as the logarithm of the intensity of the stimulus light.

Since in-phase stimulation yielded a reliably higher, and out-of-phase stimulation a reliably lower CFF than that obtained with monocular stimulation, it is obvious that the differences in CFF's for the conditions of in-phase and out-of-phase stimulation must also be statistically significant. Attention should be focused on these differences since they indicate a differential interaction of the two eyes under various conditions of binocular stimulation.

Although little is known, as yet, concerning the nature of the processes involved, the results of the present investigation are regarded as offering definite evidence for the importance of central factors in the phenomena of flicker and fusion.

Summary

The experiment on flicker and fusion described here is a repetition of one performed in 1904 by Sherrington. In the early experiment only minor shifts in CFF were found for the various conditions of binocular and mono-

cular stimulation, which Sherrington considered insignificant as evidence for interaction between the two eyes, and he concluded that the analysis of the visual phenomenon must take place in the retinas rather than in some central mechanism of the nervous system.

During the present investigation 24 subjects were tested with an electronic flicker apparatus. Four conditions of stimulation were investigated: (a) monocular flicker, dominant eye; (b) monocular flicker, nondominant eye; (c) binocular flicker, out-of-phase; and (d) binocular flicker, in-phase.

Results indicate clearly that, compared to monocular CFF, the CFF for binocular in-phase stimulation is reliably higher, while that for binocular out-of-phase stimulation is reliably lower. In view of these frequency shifts, obtained during binocular stimulation, it seems necessary to postulate some interaction between the two eyes, an interaction which seems possible only by way of some central mechanism of the brain.

(Manuscript received for immediate publication January 11, 1950)

References

1. Bartley, S. H. *Vision, a study of its basis.* N. Y.: D. Van Nostrand Co., 1941.
2. Duke-Elder, W. S. *Text-book of ophthalmology.* Vol. I. St. Louis: C. V. Mosby Company, 1938.
3. Crozier, W. J., & Wolf, E. Theory and measurement of visual mechanisms. *J. gen. Physiol.*, 1941, **24**, 639.
4. Ireland, F. H. An electronic flicker apparatus for monocular and binocular stimulation. *J. Psychol.*, 1950, **29**, 183–193.
5. Misiak, H. Practice effect on critical flicker frequency measures. *J. Gen. Psychol.*, 1948, **38**, 251–256.
6. Misiak, H. Age and sex differences in critical flicker frequency. *J. exp. Psychol.*, 1947, **37**, 318–332.
7. Sherrington, C. On binocular flicker and the correlation of activity of "corresponding" retinal points. *Brit. J. Psychol.*, 1904, **1**, 26–60.
8. Troland, L. T. Visual phenomena and their stimulus correlations. Chap. 13 in: *A handbook of general psychology.* Worcester, Mass.: Clark Univ. Press, 1934.

AMERICAN PSYCHOLOGICAL ASSOCIATION

1515 Massachusetts Avenue, N.W.
Washington 5, D. C.

The American Psychological Association, founded in 1892 and incorporated in 1925, is the major psychological organization in the United States. With approximately 6,700 members and 2,000 affiliates, it includes most of the qualified psychologists in the country. The purpose of the APA is to advance psychology as a science, as a profession, and as a means of promoting human welfare. It attempts to further these objectives by holding annual meetings, publishing psychological journals, conducting a personnel placement service for psychologists, and working toward improved standards for psychological training and service.

In order to give recognition to the specialized interests of different psychologists, the APA includes seventeen Divisions. Any person, after becoming a member of the APA, may apply for membership in as many Divisions as he wishes. The Divisions are:

Division of General Psychology
Division on the Teaching of Psychology
Division of Experimental Psychology
Division on Evaluation and Measurement
Division on Childhood and Adolescence
Division of Personality and Social Psychology
The Society for the Psychological Study of Social Issues—a Division of the American Psychological Association
Division on Esthetics
Division of Clinical and Abnormal Psychology
Division of Consulting Psychology
Division of Industrial and Business Psychology
Division of Educational Psychology
Division of School Psychologists
Division of Counseling and Guidance Psychologists
Division of Psychologists in Public Service
Division of Military Psychology
Division on Maturity and Old Age

Each Division has its own officers. Each meets annually at the time and place of the APA meeting. Each has its own membership requirements, which in some cases are higher or more specialized than the requirements for election to the APA.

The annual meetings of the APA are held in September. An attempt is made to choose meeting sites in such a way that attendance will be convenient for members and their guests in different sections of the country at different times. Thus, the 1950 meeting will be held on September 4 to 11 at Pennsylvania State College, and the 1951 meeting in Chicago. These meetings provide the members with an opportunity to present or hear reports of psychological research and lectures on topics of psychological interest, and to participate in formal and informal discussion groups. The research reports, lectures, and discussion groups cover varied topics of current interest within the field of psychology.

The chief governing body of the APA is the Council of Representatives. It includes representatives from each of the Divisions. A Board of Directors, composed of the four officers of the APA (President, President-elect, Recording Secretary, and Treasurer) and six Council members elected by the Council, is the administrative agent of the Council and exercises general supervision over the affairs of the Association.

The APA maintains an office in Washington which acts as a coordinating center for all APA activities. Journal subscriptions, membership applications, personnel placement requests, and general APA business procedures are handled in this office.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The issues listed below are available for sale

Year	Volume	Available Numbers						Price per Number	Price per Volume
1916	1	1	2	3	4	5	6	$1.25	$7.00
1917	2	1	2	–	4	5	–	$1.25	$5.00
1920	3	–	2	3	4	5	6	$1.25	$6.25
1921	4	1	2	3	4	5	6	$1.25	$7.00
1922	5	–	–	3	4	5	6	$1.25	$5.00
1923	6	–	–	–	–	–	–	——	——
1924	7	1	2	–	–	5	–	$1.25	$3.75
1925	8	1	2	3	4	–	6	$1.25	$6.25
1926	9	–	2	3	–	5	6	$1.25	$5.00
1927	10	–	–	3	4	5	6	$1.25	$5.00
1928	11	1	2	3	4	5	6	$1.25	$7.00
1929	12	1	2	3	4	5	6	$1.25	$7.00
1930	13	1	2	3	4	5	6	$1.25	$7.00
1931	14	1	2	3	4	5	6	$1.25	$7.00
1932	15	1	2	3	4	5	6	$1.25	$7.00
1933	16	1	2	3	4	5	6	$1.25	$7.00
1934	17	1	2	3	4	5	6	$1.25	$7.00
1935	18	1	2	3	4	5	6	$1.25	$7.00
1936	19	1	2	3	4	5	6	$1.25	$7.00
1937	20	1	2	3	4	5	6	$1.25	$7.00
	21	1	2	3	4	5	6	$1.25	$7.00
1938	22	1	2	3	4	5	6	$1.25	$7.00
	23	1	2	3	4	5	6	$1.25	$7.00
1939	24	1	2	3	4	5	6	$1.25	$7.00
	25	1	2	3	4	5	6	$1.25	$7.00
1940	26	1	2	3	4	5	6	$1.25	$7.00
	27	1	2	3	4	5	6	$1.25	$7.00
1941	28	1	2	3	4	5	6	$1.25	$7.00
	29	1	2	3	4	5	6	$1.25	$7.00
1942	30	1	2	3	4	5	6	$1.25	$7.00
	31	1	2	3	4	5	6	$1.25	$7.00
1943	32	1	2	3	4	5	6	$1.25	$7.00
	33	1	2	3	4	5	6	$1.25	$7.00
1944	34	1	2	3	4	5	6	$1.25	$7.00
1945	35	1	2	3	4	5	6	$1.25	$7.00
1946	36	1	2	3	4	5	6	$1.25	$7.00
1947	37	1	2	3	4	5	6	$1.25	$7.00
1948	38	1	2	3	4	5	6	$1.25	$7.00
1949	39	1	2	3	4	5	6	$1.25	$7.00
1950	40	By subscription, $7.00						$1.25	——
								List price, Volumes 1 through 38	$253.25
								30% discount	75.98
								Net price, Volumes 1 through 38	$177.27

Information about the Journal of Experimental Psychology: as noted in the table, the journal was not published in 1918 and 1919; from 1937 to 1943 it was published at the rate of two volumes per year. Nineteen numbers affecting eight volumes are out of print in the entire series.

Information about prices: when all numbers are available, the price per volume throughout the series is $7.00. For incomplete volumes, the price is $1.25 for each available number. For foreign postage, $.25 per volume should be added. The American Psychological Association gives a 30 per cent discount on orders of $150.00 and over.

Address orders and subscriptions to

AMERICAN PSYCHOLOGICAL ASSOCIATION

1515 Massachusetts Avenue N.W. Washington 5, D. C.